PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 12 JANUARY 15, 1926 Number 1

Library Fund

THE EFFECT OF VARYING MASS ON A BINARY SYSTEM: A CORRECTION

By Ernest W. Brown

Yale University, New Haven

Communicated December 4, 1925

In a recent number of these Proceedings,[1] I discussed a paper[2] by Dr. J. H. Jeans on the changes resulting in the orbit of a binary system from a loss of mass. Assuming that M (a certain function of the masses) was a function of the time, he obtained the result that the major axis a of the system varied inversely as M, while the variation of e, the eccentricity, depended on the form assumed for M but was zero in the special case he discussed. In my paper I showed that if we put $M = f(k\theta)$, where θ is the angle described after any time by the line joining the two bodies, and if we could neglect k^2, then $Ma(1 - e^2)$, Me would be constant. I then made an error in deducing the statement that $M = f(k\theta)$ would give substantially the same result as the assumption $M = f(knt)$, where n is the mean angular velocity in the orbit at any time.*

In a further note[3] Dr. Jeans gives the correct equations Ma = constant, $Ma(1 - e^2)$ = constant, for M a function of kt, where k^2 may be neglected. These results have been substantially obtained also by Professor W. D. Macmillan[3] in whose paper will be found a more detailed investigation. What we therefore have is the following: If M is a function of t only, e, Ma are constant, while if M is a function of θ only, $Ma(1 - e^2)$, Me are constant. In order to explain increasing eccentricity with increasing distance observed in the statistics of binary systems, on the basis of decreasing mass, it would therefore be necessary to assume that the rate of decrease depended to some extent on the positions of the bodies in their relative orbit; an increase of the (negative) rate at periastron would provide for an increase of eccentricity, but the change of rate would have to be comparable with the whole rate in order to be effective. The observed eccentricities are, however, very small in general when the bodies are close together, so that Dr. Jeans' conclusion that a secular decrease of mass

* Since $d^2f(k\theta)/dt^2$ is not divisible by k^2.

cannot account for the high eccentricities observed in more separated systems seems to be final.

With reference to the particular law of motion which should be adopted, that is, whether we should equate the force to the product of the mass and the acceleration or to the rate of change of momentum, it is true that the latter violates the principle of relativity. My point of view was, however, rather to find out what indications as to either hypothesis could be obtained from observation without making any assumptions as to the manner in which the mass was lost. With the corrected theory of the effect of a loss of mass, the question loses much of its interest since at present it is correlated only with a single set of observed data. But it is not impossible that certain consequences may be effective in the solar system, particularly in the critical resonance cases and in the orbits of comets.

[1] These PROCEEDINGS, **11**, No. 5, May, 1925.
[2] *Monthly Notices R. A. S.*, **85**, No. 1, Nov., 1924.
[3] *Ibid.*, No. 9, Oct., 1925.

THE DIRECT RESULTS OF MENDELIAN SEGREGATION

BY CHARLES E. ALLEN

DEPARTMENT OF BOTANY, UNIVERSITY OF WISCONSIN

Read before the Academy November 11, 1925

The exact nature of the processes concerned in the segregating (reduction) divisions constitutes one of the major problems of biological research. The genetic study of diploid organisms affords a very indirect approach to this problem; since it is necessary to reason backward from the apparent genotypic constitution of the organism through the zygote from which it arose to the constitution of the gametes whose union formed the zygote—and, in work on seed plants, still further backward through the gametophytes to the constitution of the spores from which they arose—before the segregating division in question can be considered. Less involved is the reasoning from the characters of haploid organisms developed from spores whose nuclei were formed by the segregating divisions. The nearest possible approach at present to a direct study of the effects of chromosome segregation is afforded in certain cases in which it is possible to recognize the four spores formed by division from a single mother cell. The characters of the haploid plants resulting from the germination of the spores of such a tetrad display, in effect, the immediate results of the segregation of the chromosomes or of their parts. *Sphaerocarpos Donnellii* Aust. furnishes an opportunity for a study of this nature, since (except in one

line of descent thus far studied) the four spores derived from a single mother cell remain firmly adherent.

In a previous paper,[1] dealing with a few matings involving in each case two pairs of alternative characters, it was pointed out that a spore tetrad borne by a sporophyte produced by such a mating consists in the majority of cases of two genetic types of spores only; but that occasionally all the spores of a tetrad are genetically different. The present paper presents somewhat more extensive results of matings in which the alternative pairs of characters were respectively maleness and femaleness, polyclady and non-polyclady. Descriptions of the marked and constant differences between a polycladous and a typical race have been given by the present author[2] and by Wolfson.[3] The matings whose results are here reported are the following:

Mating 28: 20.337 (♀ typical) × 21.148 (♂ polycladous).
Mating 29: 21.85 (♀ typical) × 20.266 (♂ polycladous).
Mating 30: 21.153 (♀ typical) × 20.266 (♂ polycladous).
Mating 31: 20.205 (♀ typical) × 20.254 (♂ polycladous, tufted).
Mating 32: 20.205 (♀ typical) × 20.255 (♂ polycladous, tufted).

A brief summary of the results of matings 28 and 29 was given in table 1 of a previous paper.[1]

The female clones used in matings 28 and 30 were originally classed as doubtful (possibly atypical). Matings with typical male clones, however,[4] have shown them to be genetically indistinguishable from typical females; they are consequently here considered as typical.

The male clones used in matings 31 and 32 were among the polycladous offspring of a cross (mating 14) of ♀ tufted × ♂ polycladous.[2] Among the offspring of that cross were none that showed recognizably a combination of tufted and polycladous characters; and it was concluded that in some of the polycladous offspring the extreme aberrancies of polyclady masked the lesser aberrancies of tuftedness. That conclusion is confirmed by the results of matings 31 and 32; for among the offspring in both cases were tufted, polycladous, and typical clones of each sex. Both male clones here used as parents (20.254 and 20.255) are, therefore, genetically tufted as well as polycladous.

For the present the distribution of the polycladous and non-polycladous but not that of the tufted and non-tufted characters in these two families will be considered. The distinction between tuftedness and non-tuftedness is more difficult and more subject to error than is that between polyclady and non-polyclady; and this phase of the problem is therefore left for more extended study and later report.

The distribution of the offspring of the spore tetrads in each of the five matings here considered, as to sex and the presence or absence of

polyclady, is shown in table 1. It will be observed that of the spores of any tetrad, one, two, three, or all germinated; and that, of the clones obtained from these spores, some died before their characters could be satisfactorily determined. Only those clones are tabulated whose character was determined with reasonable confidence.

TABLE 1

RESULTS OF GERMINATION OF SPORE TETRADS FROM SPOROPHYTES PRODUCED BY MATINGS OF NON-POLYCLADOUS FEMALES × POLYCLADOUS MALES

(p. = polycladous; non-p. = non-polycladous)

CLASSES ♀	♂	NOS. OF TETRADS IN MATINGS 28	29	30	31	32	TOTALS	
2 non-p.,	2 p.	..	..	..	4	2	6	
2 non-p.,	1 p.	..	..	..	3	4	7	
2 non-p.		1	1	..	3	3	8	
1 non-p.,	2 p.	..	1	..	2	1	4	
1 non-p.,	1 p.	1	..	..	2	2	5	
1 non-p.		2	2	..	4	3	11	
	2 p.	1	3	..	..	3	7	
	1 p.	..	..	2	2	..	4	52
2 p.,	2 non-p.	..	..	..	2	..	2	
2 p.,	1 non-p.	..	..	..	..	3	3	
2 p.		..	..	..	..	1	1	
1 p.,	2 non-p.	..	1	..	1	1	3	
1 p.,	1 non-p.	1	3	..	4	3	11	
1 p.		4	1	..	1	4	10	
	2 non-p.	1	1	..	4	1	7	
	1 non-p.	2	2	..	2	1	7	44
1 non-p., 1 p.,	1 non-p., 1 p.	1	..	..	..	1	2	
1 non-p., 1 p.		..	..	..	1	..	1	
1 non-p.,	1 non-p., 1 p.	..	..	..	1	..	1	
1 non-p.,	1 non-p.	..	..	..	1	..	1	
1 p.,	1 non-p., 1 p.	..	..	..	..	2	2	
1 p.,	1 p.	..	..	..	..	1	1	8
		14	15	2	37	36		104

The results summarized in table 1 agree with the expectation that in each tetrad two spores will be genetically male and two genetically female; and also that, following a mating involving these characters, two spores of each tetrad will be genetically polycladous and two genetically non-polycladous. In this respect the results are in harmony with previous experience; and in what follows it will be assumed that such a distribution

of hereditary potentialities is the general rule. If exceptions to this method of distribution occur, their occurrence is too rare to affect such statistical conclusions as will be ventured in the present discussion.

Of the tetrads the results of the germination of some or all of whose spores are shown in table 1, 52 may have been (in terms of hereditary potentialities) composed each of two female non-polycladous and two male polycladous spores; and 44 may have been composed each of two female polycladous and two male non-polycladous spores. Eight tetrads evidently showed a different type of distribution, involving (clearly in two cases) the appearance in each tetrad of four types of spores: one female non-polycladous, one female polycladous, one male non-polycladous, and one male polycladous. The results therefore indicate that in the majority of the cases here considered the segregation in the divisions that formed the spores of a tetrad was such as to produce only two combinations of hereditary potentialities and hence only two types of spores; but that in a minority of instances segregation was such as to result in tetrads consisting each of four types of spores.

This conclusion may be tested by a further analysis of the results reported in table 1. Assuming that two spores of each tetrad are (genetically) female and two male, and that two are non-polycladous and two polycladous, if the genetic constitutions of only three spores are known from the observation of the plants produced by their germination, the constitution of the fourth spore of the tetrad can be stated with confidence. Likewise, if only two spores of a tetrad germinated and the resultant plants were both female or both male or both non-polycladous or both polycladous, it is possible to conclude as to the constitutions of the other two spores. On the other hand, if only one spore of a tetrad germinated, or if the germination of but two resulted in the appearance of a non-polycladous female and a polycladous male, or of a polycladous female and

TABLE 2

GENETIC TYPES OF SPORE TETRADS (DECISIVE CASES ONLY), OFFSPRING OF NON-POLYCLADOUS FEMALES × POLYCLADOUS MALES

	2 ♀ NON-P., 2 ♂ P.	2 ♀ P., 2 ♂ NON-P.	1 ♀ NON-P., 1 ♀ P., 1 ♂ NON-P., 1 ♂ P.	TOTALS
Mating 28	5	2	0	7
Mating 29	2	1	1	4
Mating 31	12	7	3	22
Mating 32	13	6	4	23
Totals	32	16	8	56

a non-polycladous male, the distribution of potentialities as between the other spores of the tetrad remains uncertain. These latter classes of

results among those given in table 1 must, therefore, be omitted if only the decisive cases are to be considered.

Table 2 shows the distribution of the tetrads considered in table 1, after eliminating the uncertain classes. They fall into three groups, of which the first and second consist of tetrads of two spore-types, and the third includes tetrads composed of all four possible spore-types.

Table 2 again shows that, in each family as well as in the population as a whole, what may be called "four-type" tetrads were in a decided minority—this class constituting but one-seventh of the whole population. The number of cases is still too small to justify a discussion of numerical proportions. It is evident, however, that the type of segregation which produces four-type tetrads occurs, and that it occurs in a minority of cases.

The bearing of these conclusions upon the question of the nature of the reduction divisions has already been discussed.[1] In that discussion it was noted that Wettstein,[5] working with Funaria, and Kniep,[6] studying Aleurodiscus, obtained evidence of the occurrence of only two-type tetrads, and that their results led both authors to conclude that chromosome-segregation occurs only in the first reduction (heterotypic) division. More recently, studies by Hanna on *Coprinus lagopus*, which bears spores considered, from the sexual reactions of the mycelia developed from them, to be of four kinds, have been reported by Buller[7] and by Hanna himself.[8] Of 13 tetrads whose germination is reported by Hanna, each of 7 consisted of two types of spores and each of 6 of four types. Both Buller and Hanna conclude that the occurrence of chromosome-segregation in the second reduction (homoeotypic) division is thus demonstrated. As I have previously pointed out, the production in a minority of cases of four-type tetrads does seem to show that the homoeotypic division plays a part in the segregation of chromosomes or of parts of chromosomes; but it does not demonstrate that such segregation occurs *exclusively* in the homoeotypic division.

In this connection should perhaps be cited Pascher's[9] observation, in one of 13 cultures, each derived from a single heterozygote of Chlamydomonas, of offspring of four types.

Of the two types of two-spore tetrads, as shown in table 2, that composed of spores representing the character-combinations of the parental gametophytes (female non-polycladous and male polycladous) predominates over the type representing recombinations of parental potentialities (female polycladous and male non-polycladous) by a ratio of 2 to 1. The figures in table 1, however, show that the tetrads *possibly* belonging to the former class exceed those *possibly* of the latter class in a much smaller proportion (52:44). The apparently significant discrepancy in numbers between the two classes of two-type tetrads suggests the occurrence of some sort of linkage between the sexual and vegetative characters here considered.

[1] Allen, C. E., *Proc. Amer. Phil. Soc.*, **63**, 1924 (222–235).
[2] Allen, C. E., *Genetics*, **10**, 1925 (1–16).
[3] Wolfson, A. M., *Amer. Jour. Bot.*, **12**, 1925 (319–326).
[4] Allen, C. E., *Genetics* (in press).
[5] Wettstein, F. v., *Biol. Zentralbl.*, **43**, 1923 (71–83); *Zeitschr. Ind. Abst. Vererb.*, **33**, 1924 (1–236).
[6] Kniep, H., *Verh. Phys.-med. Ges. Würzburg, N. F.*, **47**, 1922 (1–29).
[7] Buller, A. H. R., *Nature*, **114**, 1924 (826).
[8] Hanna, W. F., *Annals Bot.*, **39**, 1925 (431–457).
[9] Pascher, A., *Ber. Deutsch. Bot. Ges.*, **34**, 1916 (228–242); *Ibid.*, **36**, 1918 (163–168).

ON THE ATTACHMENT OF NON-HOMOLOGOUS CHROMOSOMES AT THE REDUCTION DIVISION IN CERTAIN 25-CHROMOSOME DATURAS

BY JOHN BELLING AND A. F. BLAKESLEE

CARNEGIE INSTITUTION OF WASHINGTON

Communicated November 17, 1925

There are now at least three classes of 25-chromosome Daturas which have been more or less thoroughly studied, both by breeding experiments and by microscopical examination of their chromosomes. Of these, the main class, the *primaries*, come directly from triploids, or from non-disjunction in diploids; they give normally only diploids and primaries like themselves in their progenies; the different groupings of the chromosomes in their trivalents (Fig. 1) are those which would result in random assortment of three homologous chromosomes; they show triploid in-

FIGURE 1
Three trivalents of the primary $2n + 1$ mutant, *Rl*, with the extra chromosome in set I. The configurations are: (1) chain of three, (2) open V, and (3) ring and rod. These were drawn with the camera from iron-acetocarmine preparations.

FIGURE 2
Three trivalents of the secondary $2n + 1$ mutant, *Sg*, with the altered chromosome in set I. The configurations are: (1) closed V, (2) ring and rod, and (3) another closed V.

heritance of genes in the trivalents, in the two or more cases which have been extensively investigated; non-disjunction is apparently no more abun-

dant than in diploids; and non-homologous chromosomes have not been found attached to one another at the maturation divisions.

The second class of 25-chromosome Daturas, the *secondaries*, occur, as exceptions, in the offspring of the corresponding primaries (and of diploids and, of course, other primaries), they give normal diploids, secondaries like themselves, and the corresponding primaries, in their progeny; the groupings of the chromosomes in their trivalents (Fig. 2) point to one chromosome of the three having two homologous ends (and consequently being probably composed of two homologous halves), so that assortment is not at random; diploid inheritance of genes in the trivalent occurred in the one case investigated;[1] non-disjunction is no commoner than in diploids; and non-homologous chromosomes have not been seen connected.[2] Notwithstanding that one of the three chromosomes of the trivalent differs from the other two, measurements show that the chromosomes of the secondaries do not differ appreciably in size from those of their primaries.

In the third class of 25-chromosome Daturas, only one, *Wy*, has been as yet sufficiently investigated, both as to its breeding and its chromosomes, to allow deductions to be drawn (although some significant data have also been obtained with regard to *Hg*). The following statements then relate, primarily at least, only to *Wy*. This 25-chromosome mutant of the third class occurs in the offspring of a primary, *Pn* (which came from a *Pn* after being crossed with the *B* race of *Datura Stramonium*,[3]); *Wy* gives in its progeny, diploids, *Wy* and *Pn;* in the groupings of the three chromosomes which belong to set IX one of the three is often (in that form of *Wy* which was examined with the microscope) attached to one of the large chromosomes of bivalent I; and thus would not give free assortment nor triploid inheritance; notwithstanding this, the form of *Wy* tested by breeding showed triploid inheritance; non-disjunction was not more abundant than in diploids (but was common in *Nb*); and, as already stated, the non-homologous chromosomes, I and IX, were frequently united.

Hence there are apparently two forms of *Wy;* one which gives triploid inheritance, and one which should give approximately diploid inheritance.

The *Wy* trivalent was, in all cases seen in metaphase, an open V (although one ring and rod was seen in late prophase), but the ring and rod trivalent has been seen in the more or less parallel mutant *Hg*, which also had M-shaped chains of 4 or 5 chromosomes, groupings not yet identified with certainty in *Wy*, though they probably do occur. The trivalents observed and measured in *Wy* were 13 cases of two of the largest chromosomes, set I, forming an unequal V with one of the *Pn* chromosomes of set IX, which latter is only half the size of chromosomes I. This small chromosome was always at one free end of the V, never in the middle. Also there were two cases of an open V of three chromosomes of set IX. Lastly one case was measured where a V consisted of two small chromo-

somes of set IX, and one large chromosome of set I at a free end of the V.

From the measurements which have been made, it can be concluded that the chromosomes of sets I and IX in *Pn*, *Rl*, and *Wy*, are of the same sizes.

A working hypothesis which may be seen to fit most or all of the facts so far discovered is that in the ancestry of the B diploids there has been segmental interchange between two non-homologous chromosomes. In plants, such segmental interchange probably occurred between two non-homologous chromosomes in the ancestry of *one species* of Stizolobium (Mucuna).[5,6] In crosses between this species and three allied species and varieties, dwarf mutants with peculiar leaves appeared in F_2, in the proportion of about 0.5 per cent.[7,8] Similarly *Wy* might appear in the F_2 of crosses of normal and B diploids. The F_1 plants of the Stizolobium crosses were semisterile, as were also the F_1 plants of the crosses of normal and B diploids in Datura.

The following are the chief points that must be included in any working hypothesis:

(1) The tertiary mutants *Wy* and *Hg* have 25 chromosomes, as counted in the second metaphase.

(2) They show commonly 11 bivalents and one trivalent at the first metaphase.

(3) Non-homologous chromosomes in two particular sets (I and IX in *Wy*) are often or usually combined into trivalents (quadrivalents or quinquevalents). (See Figs. 3 and 4.)

FIGURE 3

Three trivalents from the tertiary $2n + 1$ mutant, *Wy*. The large chromosomes are from set I, while the small one comes from set IX, and is the altered chromosome (9).

FIGURE 4

Three configurations from the tertiary $2n + 1$ mutant, *Hg*. The large chromosomes belong to set I, and the smaller ones (shown in outline) perhaps to set VI.

(4) In such trivalents the two chromosomes of one set are always together (that is, not separated by a non-homologous chromosome).

(5) Thus the single chromosome of the other set is connected by one end only with the non-homologous bivalent. (Fig. 3.)

(6) Sometimes the three chromosomes of the one set (IX in *Wy*) form a trivalent.

(7) Only one chromosome of the three is ever connected with non-homologous chromosomes.

(8) The open V is the common configuration of the trivalents, but the ring and rod also occurs.

(9) The triangle or closed V does not occur.

(10) The three chromosomes (in set IX) are equal in size to the three chromosomes of the primary *Pn* (set IX).

(11) There is no perceptible difference in size or shape between the three chromosomes of set IX in *Wy*.

(12) Thus one of the chromosomes of set IX, in *Wy*, has one end normal, and attracting its fellow of set IX; while the other end is changed so that it attracts one of the ends of chromosome I.

(13) This could have been brought about by segmental interchange, at the constriction of chromosome IX; one-half of this chromosome interchanging with an equal length of chromosome I, in the ancestry of the B diploids.

In addition to these facts, the following data from the pedigreed cultures have to be reckoned with by any hypothesis.

(1) *Pn* may give *Wy* after crossing with the B strain, but without such crossing does not.

(2) Some *Wy* plants throw *Pn*.

(3) Some *Wy* plants give triploid ratios for genes in the chromosomes of set IX.

(4) *Pn* crossed twice by B diploids may give ratios, for genes in the trivalents, which show increase in the number of recessives.

(5) *Rl* after crossing with B diploids, may give ratios for genes in the IX bivalent, which are different to the ordinary diploid ratios.

(6) Crosses between the normal and B strains have given F_1 plants with 50 per cent of the pollen and ovules aborted.

A detailed working out of the hypothesis has shown that both the chromosomal and the cultural data fit the hypothesis tolerably well in most points, though there remain a few seeming exceptions (see note at end of paper). Since this hypothesis fits these facts, it may well be employed until a better is found.

Of the other $2n + 1$ mutants which apparently more or less resemble *Wy*, in the configurations of their chromosomes, viz., *Hg*, *Nb*, *Ph*, *Dv* and *X*; *Hg* has been partially studied (Fig. 4). Its largest chromosomes are larger than No. I. The meaning of this is as yet unknown.

NOTE.—All our primaries as well as the majority of our secondaries have been obtained in a purple-flowered race secured from Washington, D. C., which has been designated Line 1. When these mutants have been rendered heterozygous for certain white lines (called A whites), secured from various sources, they have all thrown disomic ratios; except *Pn* which has thrown trisomic ratios, since the genes for purple and white are located in the *Pn* chromosome. When these same Line 1 mutants are rendered

heterozygous for certain other white lines (called B whites), abnormal ratios are thrown by *Rl* and by *Pn* if the latter appears to have the constitution Pp_2. The mutant *Wy* has occurred in the offspring of parents heterozygous for Line 1 and B whites; but never in the offspring of Line 1, of A whites, nor of parents heterozygous for Line 1 and A whites. Crosses between Line 1 and three different B whites have given F_1 plants with normal pollen and ovules.

[1] Blakeslee, A. F., 1924. "Distinction between Primary and Secondary Mutants," these PROCEEDINGS, **10,** 109–116.

[2] Belling, J. and A. F. Blakeslee, 1924. "The Configurations and Sizes of the Chromosomes in the Trivalents of 25-Chromosome Daturas," these PROCEEDINGS, **10,** 116–120.

[3] Blakeslee, A. F., 1924. Quoted in *Yearbook of Carnegie Institution of Washington,* No. **23,** 1924, pp. 24–27.

[4] Belling, J., 1924. Quoted in *Yearbook of Carnegie Institution of Washington,* No. **23,** 1924, pp. 28–30.

[5] Belling, J., 1914. "On the Inheritance of Semi-sterility in Certain Hybrid Plants." *Zeitsch. ind. Abst. Vererbungslehre,* **9,** Heft 5, 303–342.

[6] Belling, J., 1915. *Rept. Fla. Agr. Exp. Sta.* for 1914. Pp. 81–105.

[7] Belling, J., 1925. "A Unique Result in Certain Species Crosses." *Zeitsch. ind. Abst. Vererbungslehre.* (In press.)

[8] Belling, J., 1925. "On the Origin of Species in Flowering Plants." *Nature,* **116,** 279. (Aug. 22).

REDDISH—A FREQUENTLY "MUTATING" CHARACTER IN DROSOPHILA VIRILIS

BY M. DEMEREC

DEPARTMENT OF GENETICS, CARNEGIE INSTITUTION OF WASHINGTON, COLD SPRING HARBOR, N. Y.

Communicated November 17, 1925

The body-color character "reddish" used in the experiments described below was first found in one of five identical pair matings which were back crosses involving several autosomal characters. Half of the males and none of the females from this pair mating were reddish, indicating that reddish is a sex-linked recessive character and that the parent female was heterozygous for it. Later experiments confirmed this interpretation.

Behavior of Reddish in Crosses with Yellow.—In appearance, reddish is very similar to the sex-linked character yellow body-color which has been known for a long time in Drosophila virilis.[1] Reddish, however, can be easily distinguished from yellow by its brighter color and still better by the color of hairs and bristles which is gray on yellow flies and yellowish on reddish flies.

Because of its similarity to yellow the first cross made with reddish was the cross between reddish and yellow, for the purpose of testing the

possibility of reddish being an allelomorph of yellow. When it was found that F_1 females from such a cross were phenotypically yellow, it was assumed that reddish was a new allelomorph of yellow. Fortunately, however, an F_2 generation was bred from that cross. In this F_2 generation of 271 flies, 112 were reddish, 129 yellow and 30 were wild-type. These results made the assumption of allelomorphism very questionable. From a cross between two allelomorphs, which behave regularly, wild-type flies should not be obtained. Since there is no precedent in Drosophila, or as far as I am aware, in any other organism, for such behavior of two recessive non-allelomorphic characters, the case of reddish indicated either a novel complementary behavior of two recessive factors in heterozygous condition or an exceptional behavior of an allelomorph of yellow.

A Map of the "Left" End of the X-Chromosome.—Since the factors located in the "left" end of the X-chromosome were extensively used in the crosses considered here, the following map of that region of the chromosome will be an aid in interpreting the data.

PILOSE	SEPIA	YELLOW	SCUTE	VERMILION
0	0.18	2.18	2.88	21.88

The map of the "left" end of the X-chromosome.

Cross between Reddish and Sepia Yellow Scute Vermilion.—Since the crosses described above indicated that reddish is an allelomorph of yellow or is located in the region close to yellow, a cross was made between reddish and sepia yellow scute vermilion—involving a majority of known factors located in that region of the X-chromosome. In the course of the experiments this same cross has been repeated over and over for other purposes, thus giving a large amount of comparable data. Table 1 gives the summary of these data obtained from 31 independent experiments involving 168 F_1 females.

In these crosses again, in addition to reddish and yellow classes non-reddish and non-yellow flies were obtained. In an ordinary case they would represent cross-overs between the reddish and yellow loci. Closer analysis of the data, however, shows that in the present case they cannot be accounted for in that way. If an assumption were made that they are crossovers it would mean that reddish is located about 3.65 units from yellow, because in a total of 9259 flies 169 cross-overs were obtained.[2] The reddish locus, therefore, should be either to the left of pilose or between scute and vermilion. The first difficulty in the analysis was met in trying to determine which of these positions was held by the locus of reddish. By taking into account all classes obtained from the crosses between sepia yellow scute vermilion and reddish it can be seen that there is no place in that region of the chromosome where the reddish locus could be placed. An analysis of the data in table 1 shows this clearly. Here,

TABLE 1

GROUPING OF MALE CLASSES FROM THE CROSS $\frac{\text{se y sc v}}{\text{re}}$ ♀ × re OR se y sc v ♂, IN REGARD TO CROSSING-OVER IN DIFFERENT REGIONS, ASSUMING DIFFERENT POSITIONS FOR THE LOCUS OF REDDISH

ASSUMED ORDER OF LOCI	re AND y CLASSES										NON-re AND NON-y CLASSES		
FREQUENCY	se y sc v 3363	re 3767	y sc v 71	se re 80	se y 36	re sc v 24	se y sc 827	re v 918	y sc 3	re sc 1	+ 116	v 30	sc v 23
$\frac{\text{se y sc v}}{\text{re}}$	0[a]	0	1–2	1–2	3	3	4	4	1–2–4	3–4	1	1–4	1–3
$\frac{\text{se}\|\text{y sc v}}{\text{re}}$	0	0	1	1	3	3	4	4	1–4	3–4	1–2	1–2–4	1–2–3
$\frac{\text{se y}\|\text{sc v}}{\text{re}}$	0	0	1	1	3	3	4	4	1–4	3–4	2–3	2–3–4	2
$\frac{\text{se y sc}\|\text{v}}{\text{re}}$	0	0	1	1	2–3–4	2–3–4	4	4	1–4	2–3	3–4	3	2

assuming different positions for reddish, it is shown in what region the crossing over should have occurred to give the observed classes. From that table it can be seen that reddish does not fit at all in any of the four possible positions. In which ever position reddish is assumed to be, some of the double cross-over classes are disproportionally high. In this cross, the occurrence of any of the double cross-over classes, except those involving the fourth region, would be an exception in itself. According to the published data[4,5] and those of Dr. C. W. Metz and Miss M. S. Moses not yet published, the smallest distance in which double cross-overs have been observed is more than fifteen units in length. In this case, the distance is less than five units long. The occurrence of any double cross-over in so small a distance would be very abnormal and the occurrence of so many double cross-overs as were observed is highly improbable.

Analysis as given in table 1 does not exhaust all possible recombinations because those which would bring yellow and reddish in the same chromosome are left out. There were two reasons for doing this. The first was that those cross-over classes probably do not occur at all, as will be shown later in the discussion of the cross between pilose scute and reddish, and the second is that nothing more than a guess could be made as to the phenotype of reddish yellow flies.

Cross between Reddish and Pilose Scute.—This cross brings out more clearly the points which were obscured in the cross between reddish and sepia yellow scute vermilion because in this case all characters involved are easily distinguishable whether alone or combined. In this cross, as in the previous one, no place for reddish could be found in the part of the chromosome where reddish should be located. As may be seen from table 2 all of the three possible positions for reddish are eliminated after a comparison is made between different cross-over classes. In this cross,

TABLE 2

GROUPING OF MALE CLASSES FROM THE CROSS $\frac{\text{pl sc}}{\text{re}}$ ♀ × re OR pl sc ♂, IN REGARD TO CROSSING OVER IN THE DIFFERENT REGIONS, ASSUMING DIFFERENT POSITIONS FOR THE REDDISH LOCUS

ASSUMED ORDER OF LOCI FREQUENCY	re 1280	pl sc 1119	re sc 5	pl 9	re pl 13	sc 22	+ 62
$\frac{\text{pl sc}}{\text{re}}$	0	0	2	2	1–2	1–2	1
$\frac{\text{pl sc}}{\text{re}}$	0	0	2	2	1	1	1–2
$\frac{\text{pl sc}}{\text{re}}$	0	0	1–2	1–2	1	1	2

again, the assumed double cross-over classes are large notwithstanding the fact that the distance between the farthest loci is less than five units. They

are so large in comparison with the single cross-over classes as to make it highly improbable that reddish can be located in any of these three positions.

One of the outstanding features of this cross is the failure to obtain the reciprocal class for the wild-type class which is the largest of the recombination classes. The reciprocal class would be pilose reddish scute. By crossing pilose reddish with scute flies it was possible to obtain pilose reddish scute flies which eliminated the possibility that that combination is not viable. The other possibility—i.e., that of low viability of pilose reddish scute flies—is also eliminated, because in the cross between pilose reddish scute and sepia yellow scute vermilion 116 flies of first parental class and 100 flies of second parental class were obtained, which shows that pilose reddish scute flies are at least as viable as sepia yellow scute vermilion flies, which are considered to have fairly good viability.

The Hypothesis.—When the results obtained from the cross between reddish and yellow were discussed, two possible interpretations were suggested, i.e., (1) that reddish is not an allelomorph of yellow, but only behaves uniquely in the crosses with yellow giving yellow F_1 generation or (2) that reddish is an allelomorph of yellow, which behaves exceptionally because in F_2 of the cross, reddish $\times$ yellow, wild type flies are found.

The first hypothesis appears to be untenable, because if reddish is not an allelomorph of yellow, but is an independent sex-linked character, it would be expected that it would behave in crosses as other mendelian characters do. The results from the crosses between reddish and sepia yellow scute vermilion, and between reddish and pilose scute, show clearly that that is not the case. The F_2 data from these crosses were so exceptional that it was not possible to "locate" reddish in the chromosome in the usual manner.

On the second hypothesis (i.e., that reddish is an allelomorph of yellow which behaves exceptionally) it is necessary to find an explanation for the occurrence of wild-type flies observed in the F_2 generation of crosses between reddish and yellow. Closer examination of the data from the cross between reddish and sepia yellow scute vermilion (table 1) shows that 116 out of 169 wild-type[6] males had the reddish chromosome from which reddish has been removed, and the rest of them had the same modified chromosome which had crossed-over in the yellow scute or scute vermilion region with the sepia yellow scute vermilion chromosome. Apparently all wild-type males obtained in this cross had the reddish chromosome in which the wild allelomorph was substituted for the reddish. Crossing over is the usual way by which one factor is substituted in a chromosome for the other one. In this case, however, as already indicated, crossing over cannot explain this substitution of wild for reddish. The mechanism by means of which this substitution is made is not known as yet and the word

"mutation," used in its broadest sense, will be applied in describing it. It is assumed that reddish frequently "mutates" to wild-type.

Test of the Hypothesis.—The results of all experiments made so far, can be explained by the assumption that reddish is an allelomorph of yellow and that it mutates frequently to wild-type. When that assumption is made it is to be expected that the F_1 generation from a cross between reddish and yellow will be yellow or nearly so, and also some wild-type flies will be expected in the F_2 from the same cross. This hypothesis, also accounts for all irregularities observed in the results obtained from crosses between reddish and sepia yellow scute vermilion and between reddish and pilose scute. All wild-type classes obtained in the two crosses are results of mutation of reddish to wild-type. That separates them from the regular cross-over classes and eliminates all irregularities in behavior encountered otherwise in analysis of results.

[1] Metz, C. W., *Genetics*, **1**, 1916 (591–607).

[2] This would represent only half of the total cross-overs because no reciprocal cross-over class is represented. This feature is discussed in later paragraphs.

[3] In this and in the following table "0" signifies non-cross-over classes, "1" classes resulting from crossing over in the first region, etc.

[4] Metz, C. W., M. S. Moses and E. D. Mason, *Washington, Carnegie Institute Publ.*, **328**, 1923.

[5] Weinstein, A., these PROCEEDINGS, **6**, 1920 (623–639).

[6] Wild-type in this case refers to the body color only.

THE EXPLANATION OF HYBRID VIGOR

BY W. E. CASTLE

Communicated December 10, 1925

It is a fact well known to breeders of animals and plants that crossing two pure breeding but different strains or varieties of the same species, as a rule produces offspring more vigorous in growth than either parental variety. Many investigators ascribe this increase in vigor to the fact that the cross-breds are *heterozygous* for all genetic factors in which the parents differed. Accordingly they designate the phenomenon heterozygosis, or in the simplified terminology suggested by Shull, *heterosis*. Shull (1911) referring to some of his earlier observations on the effects of cross-breeding maize says "I assumed that the vigor in such cases is due to the presence of heterozygous elements in the hybrids, and that the degree of vigor is correlated with the number of characters in respect to which the hybrids are heterozygous." . . . . "Mr. A. B. Bruce proposes a slightly different hypothesis in which the degree of vigor is assumed to depend upon

the number of *dominant* elements present rather than the number of *heterozygous* elements."

Shull here outlined two alternative explanations of hybrid vigor, one assigning it to heterozygosis, the other to the cumulative action of dominant factors.

The first of these explanations, which Shull himself favored, was adopted by East and Hayes (1912), the second has been adopted in a modified form by Jones (1925). Jones holds that two different varieties of corn (or of any other kind of plant or animal), which though different attain about the same total amount of growth, may owe their total growth to different genetic growth factors in the two cases. Variety A may owe its growth to the cumulative action of say three genes, 1, 3, and 6; variety B, though attaining the same amount of growth may owe it to a different genetic agency, as genes 2, 4 and 5. Each of these genes is conceived as being *dominant*, so that it has equivalent effects on growth whether present in a single or in a double dose (heterozygous or homozygous). If now the two varieties are crossed, the hybrids will have greater growth energy than either pure variety, perhaps twice as great, since it will have the summated growth energy of the dominant factors contributed by both parents, (1, 3 and 6) + (2, 4 and 5).

But in the next generation (F_2), through recombination, some of the growth factors may drop out so that only exceptionally will an F_2 individual retain them all. To account further for the rarity of F_2 individuals as vigorous in growth as those of F_1, Jones supposes that the dominant factors of the respective parents, though not allelomorphs, may lie in homologous chromosome pairs, which makes them seem to repel each other in gametogenesis, and makes it difficult to recombine them in the same gamete.

Some facts have recently come to my attention which afford a test of the relative merits of the alternative hypotheses of (1) heterosis and (2) linked dominant factors for the explanation of hybrid vigor.

Japanese waltzing mice and common varieties of fancy mice are probably domestic derivatives of two different wild species (Gates, 1925). Nevertheless they can be crossed without difficulty and the hybrids are remarkably fertile, vigorous and long-lived, surpassing either parent variety in these respects. This is the usual result of crossing animals or plants which, though distant in relationship to each other, produce fertile hybrids.

Tyzzer (1915) studied the susceptibility of Japanese waltzing mice and of other mice to inoculation with a transplantable tumor which had originated spontaneously in Japanese waltzing mice. He found that Japanese waltzing mice were without exception susceptible. When they were inoculated with a bit of the tumor, it grew progressively in their bodies. All common varieties of fancy mice, which are of house-mouse origin, were found to be completely immune. If a bit of the tumor was introduced

into their bodies, it failed to grow, or if it started to grow, growth soon ceased and the tumor disappeared. But F_1 hybrids between the Japanese waltzing mouse and the house mouse were as susceptible as pure Japanese waltzing mice. Indeed the inoculated tumor grew *even faster in the F_1 hybrids than in the pure Japanese waltzing mice.* See growth charts of Little and Tyzzer (1916).[1]

Among the F_2 individuals, however, susceptibility was very *rare.* Occasionally the inoculated tumor would start to grow in an F_2 mouse, but the growth as a rule soon came to an end and the tumor disappeared. Still more rarely growth did *not* come to an end but persisted as in Japanese waltzing mice. In such cases the animal was classed as susceptible.

Little and Tyzzer made a thorough genetic study of the case, and having found all other hypotheses tried unsatisfactory, adopted the following explanation of it.

(1) Susceptibility depends on a number of independent genetic factors (located in different chromosomes, as we may suppose, on the basis of Morgan's theory). Unless all of these are present simultaneously, the individual is not susceptible. (2) All are present in the Japanese waltzing mouse, as well as in the F_1 hybrids, which get each a complete set of genetic factors from that parent. But owing to their independence of each other (location in different chromosomes) they are rarely segregated together in F_2. Thus few F_2 individuals are susceptible. Little and Tyzzer say, p. 425, "It is probable that the number (of factors) is rather large, for if susceptibility were based on from twelve to fourteen factors, as many positive (susceptible) animals would occur as we have obtained." Now since the number of chromosomes in the mouse is about 20 pairs, it would seem that a susceptible F_2 animal must be one which has received a large majority of Japanese waltzing mouse chromosomes, enough to include a complete set of susceptibility factors, if such exist in the chromosomes.

The facts observed by Little and Tyzzer, as well as their interpretation of them, are in harmony with the heterosis theory. The transplantable tumor originated in the Japanese waltzing mouse. The constitution of its cellular elements must be very similar to that of the cells of the Japanese mouse itself, if not identical with them. It is not surprising that the tumor should grow well in the Japanese mouse since that is its normal habitat, or should refuse to grow in the house mouse, an unnatural habitat. It is perhaps not surprising that the tumor should grow in F_1 hybrids between Japanese and house mice, since in such hybrids every genetic element of the Japanese mouse is represented. But it *is* surprising that growth in the F_1 hybrid occurs even faster than in pure Japanese mice which are homozygous for all the needed elements of growth.

The heterosis theory finds an explanation of the accelerated growth in the same principle which causes acceleration of *all* metabolic processes in

cross-bred individuals. The F_1 hybrids from this cross grow faster, attain a greater adult size, mature earlier, breed oftener and live longer than mice of either pure parental race. All their tissues share in the acceleration. The transplanted tumor tissue grows faster in common with the tissues of the host itself. Doubtless the same would be true of house mouse tumor tissue, if this were transplanted into the hybrids, or of any other tissues of the parent varieties were they transplantable.

On the other hand, it is difficult to reconcile these same facts with the dominant factor hypothesis. There is no evidence that the house mouse contains *any* factors which favor growth of the tumor, since the tumor will not grow in house mice. We are accordingly not justified in supposing that the dominant factors of the house mouse, acting as a group (and they certainly come into the F_1 hybrid as a group) can *of themselves* accelerate growth of the tumor. There is no ground for supposing that they can *add to* the growth energy of the set of factors derived from the Japanese parent, other than by a heterosis effect. A second set of Japanese factors (which we know by themselves, as a group, *favor* growth of the tumor) should surely favor growth *more* than a set of house mouse factors (which we know by themselves will *not permit* growth of the tumor). Yet the fact remains that growth is slower in the homozygous Japanese mouse, than in a heterozygous hybrid.

The case, it seems to me, is a fairly critical one, in favor of the heterosis theory and against the dominant factor theory. It indicates that the zygote is not a mere summation of the factors contained in two gametes, but that the cross-bred state itself is a source of metabolic energy in the zygote. Is not this one of the chief reasons why bisexual reproduction is so obviously advantageous and consequently so widespread among both animals and plants?

[1] Dr. L. C. Strong informs me that the same thing is true in crosses of different varieties of the house mouse as regards susceptibility to a tumor originated in one of the varieties. I am indebted to him for calling my attention to the phenomenon both in his own material and in that of Little and Tyzzer.

PAPERS CITED

East, E. M., and Hayes, H. K., 1912. "Heterozygosis in Evolution and in Plant Breeding," *U. S. Dept. Agr., Bureau of Plant Industry, Bull.*, **243**, 58 pp.

Gates, W. H., 1925. "The Japanese Waltzing Mouse, Its Origin and Genetics," *Proc. Nat. Acad. Sci.*, **11**, pp. 651–653.

Jones, D. F., 1925. *Genetics in Plant and Animal Improvement*, viii + 568 pp., 227 figs. John Wiley & Sons, New York.

Little, C. C., and Tyzzer, E. E., 1916. "Further Experimental Studies on the Inheritance of Susceptibility to a Transplantable Tumor of the Japanese Waltzing Mouse," *J. Med. Res.*, **33**, pp. 393–453.

Shull, G. H., 1911. "The Genotypes of Maize," *Amer. Nat.*, **45**, pp. 234–252.

Tyzzer, E. E., 1909. "A Study of Inheritance in Mice with Reference to Their Susceptibility to Transplantable Tumors," *J. Med. Res.*, **21**, p. 519.

THE DENSITY AND ATOMIC WEIGHT OF HELIUM. II

BY GREGORY PAUL BAXTER AND HOWARD WARNER STARKWEATHER

T. JEFFERSON COOLIDGE, JR. MEMORIAL LABORATORY, HARVARD UNIVERSITY

Communicated December 14, 1925

In a recent report[1] we communicated preliminary results on the density of helium, obtained with one-liter globes, by a procedure previously tested with oxygen.[2] With the hope of increasing the accuracy of the experiments new determinations have been made with two-liter globes, and in order to obtain additional information concerning the compressibility of helium at low pressures, the density at a pressure of one-half atmosphere has been measured.

The apparatus used for the handling of the helium and for the density determinations was the same as that already described, except that larger globes were used and that two shorter barometers were added for use at the lower pressure. The larger globes proved considerably more difficult to weigh than the one-liter globes and necessitated much more careful thermostating of the balance room.

After nearly every series of density determinations the gas was subjected to an additional treatment with the adsorbent dehydrated chabazite chilled with liquid air, while the chabazite was out-gassed at 550° each time after it had been used. Gas lost in the operation was replaced with material which had been purified in exactly the same way.

Except in the case of the force of gravity the same constants are used in the calculations as in the earlier papers on oxygen and helium. In both the earlier papers an erroneous value for the force of gravity as determined at the Harvard College Observatory was used. The most recent observed value is 980.398,[3] which becomes 980.399 for the Coolidge Laboratory on account of a difference in level between the point of determination and the location of our apparatus of 3.3 meters, the two points being about 1.2 kilometers apart on a nearly east and west line. Because of this difficulty the earlier values for helium are given again. Four of the twelve original results are lowered by one unit in the fifth decimal place by the correction.

THE DENSITY OF HELIUM

0° AND 760 MM., AT SEA LEVEL, LAT. 45°

SERIES	NUMBER OF TREATMENTS WITH ADSORBENT	GLOBE I 1031.51 ML.	GLOBE II 1029.30 ML.	GLOBE III 1038.48 ML.	AVERAGE
1	3	(0.17853)	(0.17850)	(0.17858)	(0.17854)
2	5	0.17840	0.17848	0.17844	0.17844
3	6	0.17844	0.17845	0.17848	0.17846
4	7	0.17848	0.17841	0.17848	0.17846
Average of Series		0.17844	0.17845	0.17847	0.17845

SERIES	NUMBER OF TREATMENTS WITH ADSORBENT	GLOBE IV 2111.0 ML.	GLOBE V 2117.6 ML.	AVERAGE
7	9	0.17853	0.17841	0.17847
8	10	0.17849	0.17851	0.17850
9	11	0.17846	0.17852	0.17849
10	11	0.17842	0.17843	0.17843
11	14	0.17846	0.17846	0.17846
12	14	0.17846	0.17845	0.17846
	Average	0.17847	0.17846	0.17847

The average of all the experiments, except the first series with one-liter globes in which slightly impure gas was used, is 0.17846, while that given in our preliminary paper was 0.17845.

THE DENSITY OF HELIUM

0° AND 380 MM. AT SEA LEVEL, LAT. 45°

SERIES	NUMBER OF TREATMENTS WITH ADSORBENT	GLOBE IV	GLOBE V	AVERAGE
13	14	(0.08931)	0.08922	0.08922
14	14	0.08924	0.08924	0.08924
15	16	0.08923	0.08923	0.08923
16	18	0.08925	0.08922	0.08924
	Average	0.08924	0.08923	0.08923

Since the average density of helium found at one-half atmosphere's pressure is exactly half that found at one atmosphere's pressure it is apparent that helium by some process of compensation obeys Boyle's Law over this range. Burt[4] has already come to the same conclusion. This method of following compressibility, because of the low density of helium, is not as satisfactory at lower pressures as with most other gases.

With the use of the density of helium, 0.17846, and that of oxygen recently obtained by us with similar apparatus, 1.42898,[5] the following table gives values for the atomic weight of helium computed on the basis of various possible values for $(PV)_0/(PV)_1$ for oxygen. The results are only slightly different from those given in the earlier paper.

$\frac{(PV)_0}{(PV)_1}$ (OXYGEN)	LIMITING VALUE OF MOLAL VOLUME IN LITERS	ATOMIC WEIGHT OF HELIUM
1.00080	22.4115	3.9996
1.00085	22.4126	3.9998
1.00090	22.4137	3.9999
1.00095	22.4149	4.0002
1.00100	22.4160	4.0004

It is apparent that on any basis the atomic weight of helium is remarkably near the value 4.000, with the uncertainty affecting the fourth decimal.

We are very greatly indebted to the Carnegie Institution of Washington, to the Milton Fund for Research in Harvard University, to the United States Navy Department, to the Bache Fund of the National Academy

of Sciences and to the Cyrus M. Warren Fund of the American Academy of Arts and Sciences for generous assistance which has made this work possible.

[1] Baxter and Starkweather, these PROCEEDINGS, **11**, 231 (1925).
[2] Baxter and Starkweather, *Ibid.*, **10**, 479 (1924).
[3] *Special Pub. No. 40, U. S. Coast and Geodetic Survey*, p. 50 (1919).
[4] Burt, *Trans. Faraday Soc.*, **6**, 19 (1910).
[5] In the report on oxygen the final value 1.42901, was based on the erroneous value for *g* in the Coolidge Laboratory.

THE NATURE OF LIGHT

BY GILBERT N. LEWIS

CHEMICAL LABORATORY, UNIVERSITY OF CALIFORNIA

The following views regarding the mode of transmission of light are of so unorthodox a character that I set them down with some reluctance; but it seems to be generally admitted that the paradox of quantum theory cannot be resolved without doing violence to one or more of our cherished common notions. Some of the consequences of the theory that I am going to advance are repugnant to common sense, yet, searching in vain for an alternative and finding no physical fact of optics or of thermodynamics in opposition to the theory, I have come to regard it as a natural and indeed inevitable extension of Einstein's principle of relativity. Let me start with two statements which seem now to be well supported by experiment, although perhaps not entirely demonstrated. They are, however, the postulates upon which the following theory is built.

1. When an emitting atom loses energy of the amount $h\nu$, a particle with the same energy, the momentum $h\nu/c$, and the mass $h\nu/c^2$ may be regarded as traveling from the atom by a definite path until it is absorbed by another atom. The particle travels in a straight line except in the immediate neighborhood of material particles (or perhaps of other light particles) by which it may be deflected or reflected. In each encounter between a particle of light and another mass the two obey the simple laws of conservation of energy, mass and momentum.

2. The phenomenon of interference does not become less marked as the intensity of light becomes feebler, and therefore we may conclude *that the emission of a single light particle from a single atom is subject to the laws of interference.* If an optical system is so arranged that the light from a certain source produces light and dark interference bands on a photographic plate, and the source is now replaced by one in which the individual atoms only rarely emit their particles of light, these particles

will appear on the plate only where the bright bands formerly were, and their number on different parts of the plate will be proportional, on the average, to the former intensity of illumination.

Anyone who accepts these two postulates is obliged to admit that the phenomenon of interference is in some measure independent of the act of transition, and perhaps exists even at times when there is no transfer of light from one atom to another; in other words, that the atom which is capable of emitting light establishes a "field" determining the probability that its particle of light will reach one point or another. This last idea was first advanced by Slater,[1] was somewhat modified by Bohr, Kramers and Slater,[2] and has recently been restated by Slater[3] in its original form. He assumes that an atom capable of emitting a certain frequency possesses a virtual oscillator, of the same frequency, which produces a virtual electromagnetic field; that this field obeys the ordinary electromagnetic laws, and that the magnitude of its virtual Poynting vector determines the probability that a particle of light will reach a given spot. In the same way Swann[4] assumes a field with an energyless Poynting vector so that "we practically constrain the quanta to follow the paths of the Poynting vector." (The suggestion by L. de Broglie[5] that the Poynting vector does not determine the path of the light particle, but only whether the particle may be absorbed at a certain point, is not consistent with our first postulate, since he assumes the conservation laws to be only statistically true.)

Now in addition to an objection mentioned by Slater, there are strong arguments against the idea that the particles of light follow the path of the Poynting vector. In the first place, the ordinary Poynting vector is calculated from the sum of *all* the electric and magnetic fields at a point, while the complete lack of interference of light from distinct sources shows that the field of a single virtual oscillator must by itself produce interference, regardless of the existence of fields from other oscillators. But even if we impose this limitation upon the definition of the virtual Poynting vector, a fatal objection remains. If we admit that a particle of light traverses the distance l between two material objects in the time l/c, and that it cannot move with a velocity greater than c, it must travel in a straight line; but the Poynting vector in an interference field gives a wavy line. Suppose that by means of mirrors we make two portions of the virtual radiation traverse one another at a slight angle, we do not know by experiment that there is absolutely no loss of time in such a process, but certainly there is no such retardation as would be found if the particle of light were forced to follow the tortuous path of the Poynting vector.

Suppose that we do not attempt to decide upon the particular path taken by the particle of light and assume only that the existence of a Poynting vector in a given part of the receiving surface determines whether the

particles may reach that part. Here again we are met by contradiction. Figure 1 shows S, a source of light; A and B, two slits; C, a point to which light may go if the slit A alone is open, but which is forbidden if both A and B are open. Now consider that a shutter opens both slits for an instant; if a particle of light were about to pass through A and be deflected in the direction of C, the virtual field passing through B and moving with the velocity of light could not meet the particle until it reached some point such as D. But by this time its direction would have been fully determined and it would proceed to C. This, however, is a forbidden point.

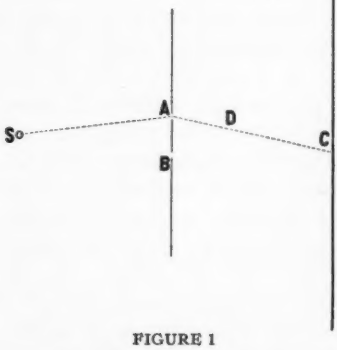


FIGURE 1

So far our analysis has merely shown that no explanation so far proposed explains at the same time the phenomena of quanta and of interference. As we study the simple system described in figure 1, there constantly recurs to us the hardly credible thought that in some manner the atom in the source S can foretell before it emits its quantum of light whether one or both of the slits A and B are going to be open. Although it sounds absurd, this, in a certain sense, is the theory that I am going to propose.

It is generally assumed that a radiating body emits light in every direction, quite regardless of whether there are near or distant objects which may ultimately absorb that light; in other words that it radiates "into space." This assumption has seemed natural and convenient. We know that on a clear night objects radiate energy into what seems empty space, but I am not aware that any exact experiments have been made at different altitudes to eliminate the effect of the atmosphere and to determine whether the emission is that which would be given by Steffan's law. In any case we do not know how much cold matter the universe may contain.

I am going to make the contrary assumption that an atom never emits light except to another atom, and to claim that it is as absurd to think of light emitted by one atom regardless of the existence of a receiving atom as it would be to think of an atom absorbing light without the existence of light to be absorbed. I propose to eliminate the idea of mere emission of light and substitute the idea of *transmission*, or a process of exchange of energy between two definite atoms or molecules. Now, if the process be regarded as a mere exchange, the law of entire equilibrium, which I have recently advanced,[6] requires us to consider the process as a perfectly symmetrical one, so that we can no longer regard one atom as an active agent

and the other as an accidental and passive recipient, but both atoms must play coördinate and symmetrical parts in the process of exchange.

I shall not attempt to conceal the conflict between these views and common sense. The light from a distant star is absorbed, let us say, by a molecule of chlorophyl which has recently been produced in a living plant. We say that the light from the star was on its way toward us a thousand years ago. What rapport can there be between the emitting source and this newly made molecule of chlorophyl? Suppose we make this same star the source of light in the apparatus of figure 1. By opening the second slit we prevent a particle of light from reaching the point C. Do we therefore prevent its original emission? If so it would mean that we could, perhaps in a trivial way, but nevertheless in principle, alter the course of past events.

Such an idea is repugnant to all of our notions of causality and temporal sequence; but we must remember that these notions have arisen from the observation of complex processes which are very different from the elementary reversible processes which we are here considering. Unless the result of some actual fact of experiment or observation can be brought against the new view we need not be deterred by this conflict with common notions. Indeed we shall see that there are already some inconsistencies between prevailing physical ideas and that geometry which so admirably interprets the kinematics of relativity.

Admitting that a radioactive substance emits α-rays and β-rays in all directions without regard to their later absorption by other matter, why should we not make the same assumption regarding light? The answer is suggested by the new geometry. Let us consider the four-dimensional manifold of relativity,[7] which for simplicity may be represented in figure 2 by a two-dimensional diagram with one axis of space, OX, and another of time, OT. In the geometry characteristic of this space-time there is a sharp distinction between all lines of the class OX and OX', which Minkowski calls space-like lines, and all time-like lines of the class of OT and OT'. Between these two classes are other lines, the singular lines such as OL and OL', which belong to neither of the other classes and bear no more resemblance to one class than to the other.

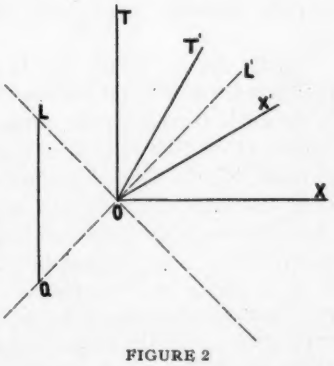


FIGURE 2

The path or locus of a material particle in time and space is always one of the time-like lines, such as OT', and the slope of this line with respect

to the chosen time axis OT represents the velocity of the particle. On the other hand, the slope of a space-like line, such as OX', does not even suggest a velocity. Now in spite of the symmetry demanded by the geometry we affiliate the line OL with the time-like lines and say that it represents the velocity of light. As a concession to traditional thought we do violence to our geometry when we call the light process a form of motion at all; and while we shall continue to make this concession, it will be with a realization of the unique character of radiation.

There is another remarkable feature of this geometry which has been insufficiently employed in physics. If two lines such as OT and QL represent the space-time loci of two material particles, then by the methods of measurement that characterize this geometry, the intercepts OL and OQ of singular lines between these loci, are of zero length. This is an idea of which much use has been made in the mathematics but none in the physics of relativity. The proposals which I am making in this paper are tantamount to assuming that such a distance is also zero in a physical sense, and that two atoms whose loci are OT and QL may be said to be in *virtual contact* at any two points such as O and L or O and Q, which are connected by singular lines. When two atoms are in ordinary physical contact we do not inquire how one atom ascertains that the other is in a position to receive energy, nor need we so inquire in the case of virtual contact, even though the time of emission and the time of absorption are said to be separated by thousands of years. Such a statement depends upon an arbitrary choice of a time axis. If in our figure we should take, not OT, but time axes lying nearer and nearer to OL, not only the time elapsing between O and L but also their spatial distance would *approach zero*.

Finally, let us remark that in a pure geometry it would surprise us to find that a true theorem becomes false when the page upon which our figure is drawn is turned upside down. A dissymmetry alien to the pure geometry of relativity has been introduced by our notion of causality. For example, it is the theory of retarded potentials that if, in figure 2, OT represents the locus of an electron and QL the locus of another electron, then the curvature of the line OT at O is determined by the character of the other electron as it is at Q (using the "forward" singular line QO) and not as it is at L (using the "backward" singular line LO). There is, however, no experimental evidence for such an assumption.

In all three of the respects that I have mentioned the theorem that I am proposing, although it still makes concession to our habitual thought, goes a long way toward bringing closer correlation between our physical and our geometrical concepts: it emphasizes the unique character of the singular lines (motion of light), it makes physical use of the theorem of zero distance along singular lines, and it does away with that distinction between

past and future which, however useful it may be in other cases, seems to have no significance in a purely reversible process.[6]

Let us represent in figure 3 the path in space-time of two atoms A and B, of which A is capable of losing, and B of gaining the energy $h\nu$. For simplicity we may consider the two atoms relatively at rest and therefore their two loci parallel. Let us now, following the suggestion of Slater's virtual oscillator, draw a helix about the line A to represent in the atom A some sort of periodic motion or change of frequency ν. Let us, however, also draw a helix of the same radius and the same period about the line B. We may speak of the phase of one helix at the point A' and compare it with the phase of the other helix at the point B'. If $A'B'$ is a forward singular line and $A''B''$ is a parallel singular line, then the phase difference between A' and B' will be the same as that between A'' and B'', and may be called simply the phase difference between the two atoms.

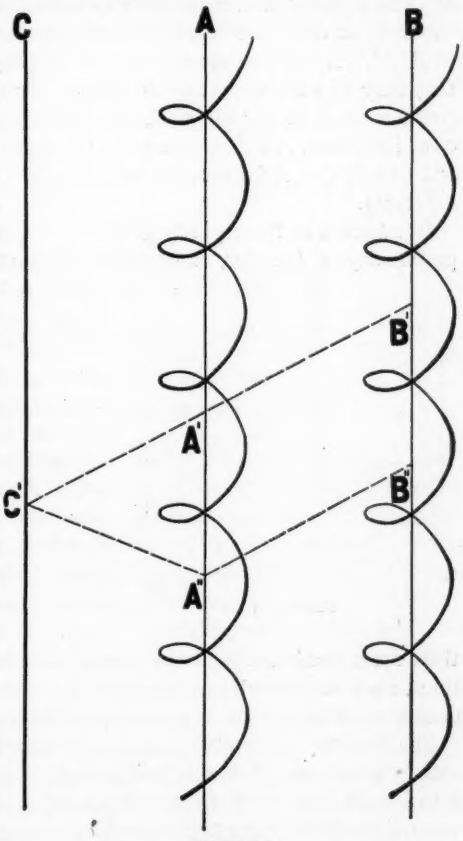


FIGURE 3

In our four-dimensional manifold the two parallel lines A and B determine a perpendicular 3-space, and if we choose a plane in this 3-space the projections of the two helices will be ellipses; but for each helix there is one plane upon which the projection is a circle. The angle between this characteristic plane of A and the characteristic plane of B may be called the relative polarization of the two atoms. Thus we speak not of the phase and polarization of light but of the relative phase and relative polarization of the atoms themselves; and we shall assume that the chance that the atom B accepts radiation from A depends upon these two quantities, and upon

the distance between A and B. I must reserve for another communication the demonstration that by proceeding in this simple geometric manner we may derive the quantitative laws of optics. But there are one or two qualitative points which should be mentioned here.

If in addition to our two atoms we introduce a third object which will serve as a mirror, and represent its locus by the line C, then we may draw the singular lines $A''C'$ and $C'B'$ to represent an alternative optical path to $A''B''$. If the two atoms were not in phase with respect to one path, they may be with respect to the other. Or again the phases of A as projected upon B by the two paths may differ by just half a period, in which case, irrespective of the phase of B, the probability of a transfer of energy will be zero (provided that the lengths of the two paths are not materially different).

In addition to these purely geometrical conditions which determine the probability of transfer, it might be supposed that a complete analysis would require knowledge of the atoms from which A obtained its energy and other atoms to which B will ultimately give its energy. But this and other complications may be avoided if we assume that the energy is held by the atom A before the transfer and in the atom B after the transfer for periods which are long compared with the time in which the particles of light may be held by the atoms of the mirror, lenses and other portions of our optical system. If this condition is not met, there are not only mathematical complications but also physical complications, such as anomalous dispersion and lack of coherence; but these are questions which we must discuss in a later paper.

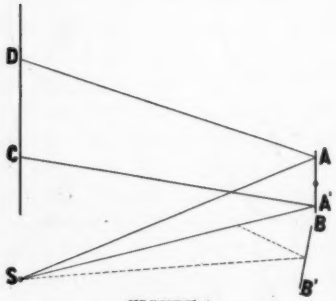


FIGURE 4

Finally, it is legitimate to inquire whether a theory which claims to resolve a serious paradox and also to bring our physical concepts more nearly into accord with the geometry of relativity, can also make a new experimental prediction. In fact there is a crucial experiment which does not seem absolutely beyond the reach of our present experimental technique. Figure 3 shows a source of light, S, and two mirrors, AA' and BB', so adjusted as to produce an interference pattern upon the plate CD. Now let us suppose that the mirror AA' is so narrow that its width is only one-half the distance between the adjoining dark and light bands C and D, and let it be suspended in such manner as to permit rotation in the plane of the diagram. According to the classical theory, and also according to the views set forth by Slater and by Swann, the Poynting vector being

the same at A and A', the pressure of light will be uniform over this mirror. According to the new theory, particles of light will pass over the path SAD where D is in the bright band, but will not pass over the path $SA'C$ if C is in the dark band. Therefore, the light pressure upon this mirror will all be on the side A and a torsion of the mirror will result.

[1] Slater, *Nature*, **113**, 307 (1924).
[2] Bohr, Kramers and Slater, *Phil. Mag.*, **47**, 785 (1924).
[3] Slater, *Nature*, **116**, 127 (1925).
[4] Swann, *Science*, **61**, 425 (1925).
[5] L. de Broglie, *Phil. Mag.*, **47**, 446 (1924).
[6] Lewis, these PROCEEDINGS, **11**, 179 (1925); **11**, 422 (1925).
[7] Wilson and Lewis, *Proc. Amer. Acad.*, **48**, 389 (1912).

THE RÔLE OF THE FARADAY CYLINDER IN THE MEASUREMENT OF ELECTRON CURRENTS

BY ERNEST ORLANDO LAWRENCE[1]

SLOANE LABORATORY, YALE UNIVERSITY

Communicated December 11, 1925

Despite knowledge of the magnitude of reflection and secondary emission of electrons, Faraday cylinders have customarily been regarded as complete absorbers of electrons. For example Lehmann and Osgood[2] recently have brought forward evidence indicating that electrons passing through small apertures are not of homogeneous velocities—this conclusion resting on the assumption that the Faraday cylinder used in their experiments absorbed at least 99% of the impinging electron stream. An exception to this usual regard of the matter is recorded in some experiments of Professor J. T. Tate[3] who found that the most efficient Faraday cylinder he devised had an absorption coefficient of about 0.95 for relatively slow velocity electrons. In view of the very important relation of Lehmann and Osgood's conclusions to much other experimental work and, indeed, because of the uncertain yet crucial status of the Faraday cylinder itself, this preliminary experimental investigation was carried out.

Electrons were accelerated through a distance of 4 cms. from a tungsten filament to a plane anode through which extended a tube 2 mm. in diameter and 16 mm. in length, the tube end nearer the filament being in the anode plane. The electrons emerging from the tube impinged on a Faraday cylinder 2.2 cms. in diameter through an opening 1 cm. in diameter at a distance of 1 mm. from the end of the tube. The distance of the closed end of the Faraday cylinder—it's effective length—was made variable by a stopcock swivel arrangement. Since the absorbing power of the Fara-

day cylinder is a function of its length, measurement of the currents to the cylinder with various retarding potentials and cylinder lengths with this arrangement gave data indicative of the percentage of the impinging electrons absorbed.

Figures 1 and 2 are typical of the experimental data obtained. The ordinates of the curves are the currents to the Faraday cylinder corresponding to retarding potentials represented by the abscissae. Curve no. 1 is for the shortest cylinder length and succeeding numbers denote data obtained with longer cylinders. As is evident, the data of figure 1 correspond to an accelerating voltage between the filament and anode of about twenty-four volts while figure 2 is representative of an accelerating voltage of approximately 100 volts. It is seen that for the lower voltage electrons there is a quite regular increase of the electron currents to the Faraday cylinder as its length is made greater while for the higher voltage electrons there occur quite drastic alterations of the current-retarding potential curves with alterations of cylinder dimensions. These latter effects are to be explained on the basis of knowledge that emission of secondary electrons is greater for the higher velocity electrons. Take curve 2 of figure 2 for example. With a zero retarding potential the electrons impinge on the cylinder generating secondary electrons some of which get out of the cylinder so that the resultant current measured is smaller than the actual initial current entering the cylinder. Application of a small retarding potential increases the proportion of the secondary electrons emerging from the cylinder and at the same time reduces the number of secondaries generated because of the decrease in the velocities of the generating electrons. It is evident for the situation represented by curve 2 the former effect is greater than the latter up to retarding potentials of about 40 volts and between 40 and 60 volts the number of secondary electrons produced decreased markedly, thus enhancing the observed current. This general circumstance persists for all cylinder dimensions although, of course, as the absorbing efficiency of the cylinder increases

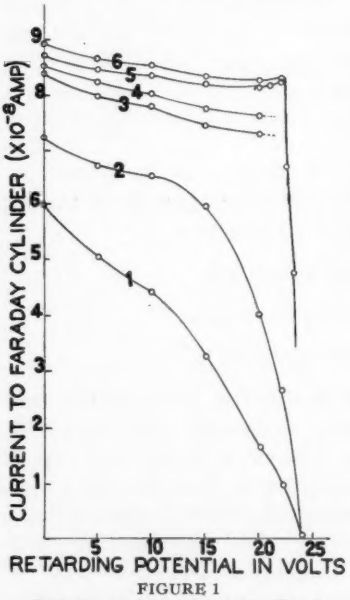


FIGURE 1
Cylinder lengths same as in Fig. 2.

these effects become less marked. Further elucidation of the data is hardly necessary. It is sufficient to emphasize the evident fact that *Faraday cylinders of quite usual dimensions do not retain all the entering electrons and current retarding potential measurements generally do not indicate accurately the distribution of velocities in an electron stream.*

These experimental results, together with other data not exhibited here, indicate that the measured currents to the Faraday cylinder are a function of several variables including the distribution of velocities of the electrons entering the cylinder as well as its absorbing efficiency. It has become clearly recognized that only by rather indirect evidence and inference is it possible to distinguish between the several factors in a manner employed in the present work. To clear up the questions here involved and, indeed, to establish on a firm, experimental basis many other doubtful facts in experimental physics an outstanding need is a beam of electrons moving with homogeneous velocities. With this in mind a method of magnetic analysis at the present moment is being developed designed to produce a beam of electrons of the order of ten volts' velocity with a velocity variation throughout the beam of about 0.1 volt.

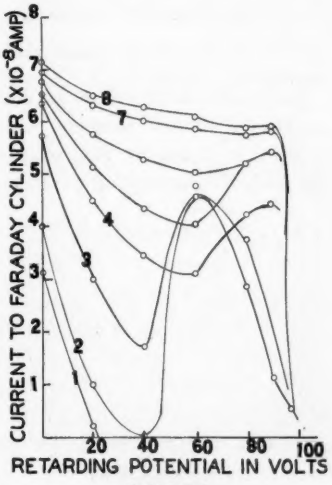


FIGURE 2
Cylinder lengths:

Curve 1, for cylinder length of 0.3 cm.
" 2, " " " " 0.8 "
" 3, " " " " 1.9 "
" 4, " " " " 3.0 "
" 5, " " " " 4.6 "
" 6, " " " " 5.7 "
" 7, " " " " 7.3 "
" 8, " " " " 9.5 "

It is a pleasure indeed to acknowledge my indebtedness to Professor W. F. G. Swann for his interest in this work.

[1] NATIONAL RESEARCH FELLOW.

[2] Lehmann and Osgood, *Proc. Cam. Phil. Soc.*, **22**, 731.

[3] Tate, *Physic Rev.*, **17**, p. 394.

THE QUANTUM THEORY OF THE DIELECTRIC CONSTANT OF HYDROGEN CHLORIDE AND SIMILAR GASES

BY LINUS PAULING*

GATES CHEMICAL LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated December 14, 1925

Under the influence of an electric field a gas whose molecules have a permanent electric moment and in addition can have a further moment induced in them by deformation becomes polarized in the direction of the field, the amount of polarization per unit volume being[1]

$$P = \frac{3}{4\pi}\frac{\epsilon - 1}{\epsilon + 2} F = N\mu \overline{\cos\theta} + N\alpha F. \qquad (1)$$

Here ϵ is the dielectric constant of the gas, F the strength of the applied field, N the number of molecules in unit volumes, μ the permanent electric moment of a molecule, and α the coefficient of induced polarization of a molecule; $\overline{\cos\theta}$ is the average value of $\overline{\cos\theta}$ for all molecules in the gas, and $\overline{\cos\theta}$ is the time-average of $\cos\theta$ for one molecule in a given state of motion, θ being the angle between the dipole axis and the lines of force of the applied field.

On the basis of the classical theory it was shown by Debye[2] that $\overline{\cos\theta}$ is given by the Langevin function

$$\overline{\cos\theta} = \coth\frac{\mu F}{kT} - \frac{kT}{\mu F}. \qquad (2)$$

This reduces, for values of μF small in comparison with kT, to

$$\overline{\cos\theta} = {}^1/_3\frac{\mu F}{kT}. \qquad (3)$$

This result must now, however, be replaced by one based upon the quantum theory. In this case a generally applicable expression cannot be obtained, and it is necessary to consider particular molecular models. W. Pauli, Jr.[3] has treated the diatomic dipole, which is of interest to us, and his treatment forms the basis of this discussion.

The Hamiltonian function for this dynamical problem, using polar coördinates with the polar axis in the direction of the lines of force, is

$$H = \frac{1}{2I}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right) - \mu F\cos\theta. \qquad (4)$$

In this expression p_θ and p_φ are the momenta corresponding to the coördinates θ and φ, and I is the moment of inertia of the rotator about

an axis through its center of mass and normal to the line joining the atomic nuclei. Applying the Wilson-Sommerfeld quantum conditions we write

$$\left.\begin{aligned} \oint p_\theta \, d\theta + \oint p_\varphi \, d\varphi &= mh, \\ \oint p_\varphi \, d\varphi &= 2\pi p_\varphi = nh. \end{aligned}\right\} \tag{5}$$

The time-average value of $\cos\theta$ for a given rotator characterized by the quantum numbers m and n is found to be, in case μF is small in comparison with the energy-constant W,

$$\overline{\cos\theta} = \frac{4\pi^2\mu FI}{m^2h^2}\left(\frac{3}{2}\frac{n^2}{m^2} - {}^1/_2\right). \tag{6}$$

Pauli assigned to m the values 1, 2, 3, $\ldots\infty$, and to n the values ± 1, ± 2, ± 3, $\ldots$ $\pm m$. We shall give m the values ${}^1/_2$, ${}^3/_2$, ${}^5/_2$, $\ldots\infty$, and n the value $\pm {}^1/_2$, $\pm {}^3/_2$, $\pm {}^5/_2$, $\ldots$ $\pm m$; for these half-quantum numbers for m are indicated by the infra-red oscillation-rotation absorption spectrum[4] of hydrogen chloride and definitely required by the recently observed[5] pure rotation absorption spectrum, and it can be shown that similar values for n are reasonable. Assigning equal *a priori* probabilities to all of the possible quantum states characterized by half-integral values of m and n we have for the probability of a given state

$$w(m, n) = \frac{e^{-\frac{W}{kT}}}{\sum_{m={}^1/_2}^{\infty}\sum_{n=\pm {}^1/_2}^{\pm m} e^{-\frac{W}{kT}}}. \tag{7}$$

Introducing the variable $\sigma = \dfrac{\Theta}{T} = \dfrac{h^2}{8\pi^2 IkT}$ we obtain the following expression for $\overline{\cos\theta}$:

$$\overline{\cos\theta} = \frac{\mu F}{kT}\frac{1}{2\sigma}\frac{\sum_{m={}^1/_2}^{\infty}\sum_{n=\pm {}^1/_2}^{\pm m}\frac{1}{m^2}\left(\frac{3}{2}\frac{n^2}{m^2} - {}^1/_2\right)e^{-\sigma m^2}}{\sum_{m={}^1/_2}^{\infty}\sum_{n=\pm {}^1/_2}^{\pm m} e^{-\sigma m^2}}. \tag{8}$$

Pauli, who obtained a similar expression with whole quantum numbers, evaluated the sums by assuming σ to be very small. This is not justifiable, however, in cases of practical interest, for which σ has values near 0.05, for it then leads to results in error by 30% or more. Replacing m by $r - {}^1/_2$ and n by $s - {}^1/_2$ and summing for s one obtains

$$\overline{\cos\theta} = \frac{\mu F}{kT}C, \tag{9}$$

where

$$C = \frac{1}{4\sigma} \frac{\sum_{r=1}^{\infty} \frac{r}{(r - 1/2)^3} e^{-\sigma(r-1/2)^2}}{\sum_{r=1}^{\infty} r\, e^{-\sigma(r-1/2)^2}}. \qquad (10)$$

C has been evaluated for a range of values of σ by direct substitution. The values obtained are given in table 1.

TABLE 1

σ	C	σ	C
0.20	2.848	0.04	3.690
0.10	3.264	0.03	3.796
0.07	3.445	0.02	3.929
0.05	3.599	0	4.570

In order to apply the results of the theory in the interpretation of the experimental measurements on hydrogen chloride, made by Zahn,[6] it is necessary to know Θ for this substance. Czerny[5] gives for $\frac{hc}{8\pi^2 I}$ the value 10.397, which leads to a value of 14.9° for Θ. It is found that Zahn's experimental results conform reasonably well with the theoretical curve corresponding to a value of 0.3316×10^{-18} c.g.s.u. for μ and 0.00077 for $4\pi N_0\alpha$, with N_0 the number of molecules per cc. at standard conditions. These values differ considerably from those obtained from the same data with the classical theory, 1.034×10^{-18} and 0.00104.

The experimental points, which surprisingly enough lie very close to a curve of the type required by the classical theory, show deviations as large as 2% from the best quantum theory curve. Nevertheless it is not easy to reject the straight-forward and well-grounded quantum theory calculations given in this paper; possibly the discrepancies can be attributed to errors in the experimental measurements. Further measurements testing this point are needed.

The quantum theory value for $4\pi N_0\alpha$ is confirmed by independent evidence; for it conforms with Maxwell's relation $4\pi N_0\alpha = n_0^2 - 1$, in which n_0 is the index of refraction of the gas at standard conditions for light of frequency far from any frequency characteristic of the molecule. It is known that n as determined with light of ordinary wave-length is larger than n_0; hence the observed value 0.000888 for $n^2 - 1$ for hydrogen chloride with the sodium D-lines cannot be reconciled with the classical value 0.00104 for $4\pi N_0\alpha$, but supports the quantum theory value 0.00077.

Similar calculations for hydrogen bromide lead to the values 0.252×10^{-18} for μ and 0.00102 for $4\pi N_0\alpha$. The data for hydrogen iodide are inaccurate because of dissociation, and cannot be used.

The electric moment found for hydrogen chloride is equal to that of a dipole composed of a proton and an electron 0.0694 Å apart; for hydrogen bromide the equivalent distance is 0.0528 Å.

A more extensive account of these calculations will be presented to the *Physical Review* for publication.

* NATIONAL RESEARCH FELLOW IN CHEMISTRY.

[1] See, for example, Debye in Marx "Handbuch der Radiologie," Leipzig Akademische Verlagsgesellschaft, **1925**, V. 6, p. 619.

[2] P. Debye, *Physik. Zeit.*, **13**, 97 (1912).

[3] W. Pauli, Jr., *Z. Physik*, **6**, 319 (1921).

[4] Colby, *Astrophys. J.*, **58**, 303 (1923).

[5] Czerny, *Z. Physik*, **34**, 227 (1925).

[6] Zahn, *Physic. Rev.*, **24**, 400 (1924).

THE MOBILITY OF GASEOUS IONS IN HCl GAS AND HCl AIR MIXTURES

BY L. B. LOEB

PHYSICAL LABORATORY, UNIVERSITY OF CALIFORNIA

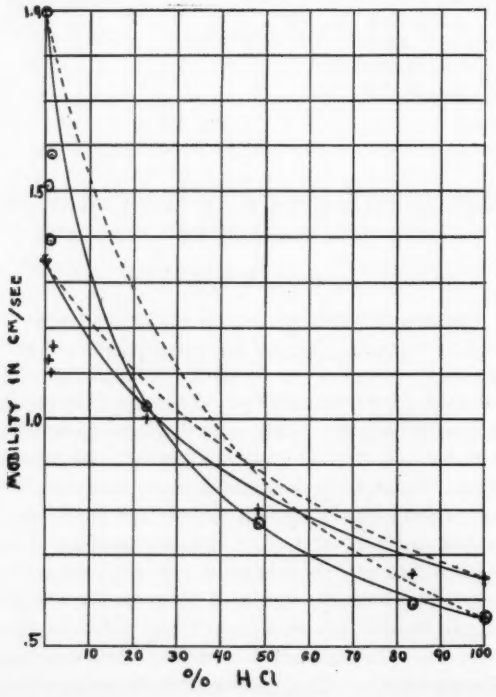
Communicated November 19, 1925

Introduction.—In a further endeavor to answer the question as to the nature of the gaseous ion the writer undertook measurements of the mobilities of positive and negative ions in mixtures of HCl gas and air of varying composition. Practically no systematic measurements of mobilities in mixtures of varying concentrations of different gases have been made. Those of Blanc[1] in mixtures of H_2, CO_2 and air and those of Loeb and Ashley[2] on mixtures of NH_3 and air constitute the only ones up to the present. Certain differences between the results of Blanc on the mixtures of the less reactive gases and those of the writer and Miss Ashley on mixtures in NH_3 and air made it seem of importance to continue the investigation with other gases of dissimilar nature. The particular differences observed for NH_3 and air mixtures and their theoretical interpretation will form the subject of another paper.

The only mobility measurements made on HCl are those made by Rutherford[3] in 1897 who found that the sum of the positive and negative mobilities in HCl was in the neighborhood of 1.27 cm./sec. The results of the measurements of mobilities in pure HCl as well as in mixtures of HCl and air yielded such curious results that it seemed of importance to publish a preliminary account of them in this article.

Experimental Procedure.—The measurements were made by the Franck[4] modification of the Rutherford alternating current method, used by

the writer[2] in previous work. In general the procedure was carried out in very much the same fashion as before. Because of the corrosive nature of the gas used a new and smaller ionization chamber was made which was plated with gold on the inside. To avoid contact potentials the gauze and plates were also gold plated. The chamber opened from the top, and a gas-tight joint was insured by screwing the top down onto a leaden gasket by means of sixteen screws. The perforated brass gauze was made plane by soldering it to a flat iron ring while hot. Parallelism between gauze and



upper plate was achieved in spite of the gasket joint, by a ball and socket union between the threaded sleeve which raised and lowered it, and the upper plate. The sleeve and upper plate could be raised and lowered relative to the gauze by means of the insulated screw supporting them and an electromagnet operated from outside. Thus the upper plate was set nearly parallel to the gauze and when the chamber was sealed it was lowered by means of the electromagnet until it had pressed itself parallel to the gauze.

The gauze had a diameter of 7.5 cm. was 0.038 cm. thick and had 25 holes to the cm.^{2} each hole being 1.1 mm. in diameter. The depth of the

auxiliary field was 1.3 cm. and the distance between gauze and the upper plate was in the neighborhood of 1 cm. The auxiliary field used throughout was 4.5 volts. The frequency of alternation was so chosen in these measurements that the the feet of the mobility curves were confined between 30 and 60 volts. This tended to eliminate the errors due to the differences between alternating and auxiliary field strengths discussed elsewhere.[5] Most measurements were made at frequencies of from 30 to 60 alternations per second and were of square wave form.

The HCl was prepared by letting Baker's C.P. H_2SO_4 drop onto Baker's C.P. NaCl which had been baked 6–8 hours at 100°C. in vacuo. The HCl generator was exhausted to a small fraction of a mm. before the H_2SO_4 was admitted. The HCl then passed through two liquid air traps cooled to −100°C. by frozen ethyl alcohol into a trap immersed in liquid air. Here it condensed to a solid. After a sufficient amount had collected the generator was shut off and the trap was pumped down to an X-ray vacuum or better by a mercury vapor pump. This removed traces of air liberated from the NaCl crystals. The HCl was then distilled over into the ionization chamber to the requisite pressure. When working in mixtures the HCl was admitted first and the air dried over $CaCl_2$, P_2O_5 and liquid air was run in afterward. For small pressures of HCl, HCl was run in up to 10 cm. pressure first, let stand a bit, and pumped down to the required pressure. Air was then admitted to bring the total pressure to 760 mm. It was found convenient to remove any excess HCl gas by means of a water aspirator pump, which efficiently kept the HCl from getting into the room.

In passing it is of interest to note that the peculiar pink modification of solid HCl was observed, which had first been noticed by Giauque and Wiebe in the chemistry department of the University of California. An account of a study of this peculiar product will shortly be published by those authors.

In the later repetition of the work a great deal of trouble was encountered with contact differences in potential. This persisted in spite of the most careful gold plating of the electrodes. It is suspected the trouble arose through the fact that the upper plate was a bronze casting while the gauze was brass. Apparently even in the plating sufficient differences in the surfaces resulted to cause the effect. This trouble with that observed in NH_3[2] indicate that great care must be used to avoid such effects where chemically reactive gases are involved. The presence of contact potentials can always be suspected when the ratio of the mobilities of positive and negative ions in air differ widely from the commonly accepted values. They also make themselves noticeable by causing a current between the plates with positive or negative ions (depending on the sign of the effect), when no field is on. By tedious control measure-

ments these may be corrected for. Fortunately, in the first set of measurements which comprised the majority of the results, these effects were sufficiently small to be neglected.

Mobilities in Pure HCl.—The mobilities found in pure dry HCl gas are listed below. The values starred have been corrected for contact potentials, the others not. All the measurements were made between 21°C. and 24°C. and at pressures between 740 and 780 mm. of Hg. They were all reduced to a pressure of 760 mm. by assuming that the mobility was inversely proportional to the pressure, an assumption which was later verified. The mobilities are in cm./sec. per volt/cm.

MOBILITIES IN PURE HCl

POSITIVE IONS	NEGATIVE IONS
0.640	0.563
0.616	0.556
0.680	0.572
0.636	0.554
0.642*	0.560*
0.656	0.558
0.662*	0.525*
0.648 Average	0.555 Average

The uncertainties due to the contact potentials do not permit one to attribute greater accuracy to the values than are represented by 0.65 and 0.56 cm./sec. per volt/cm. for the positive and negative ions, respectively.

Discussion of the Results.—The values above for the mobilities give a positive ion mobility which is unmistakably and definitely higher than that for the negative ion. This point is one of great interest. While it had previously been believed that in some gases the positive ions had greater mobilities than the negative ions[6] a careful investigation by Yen[7] showed this to be wrong in all cases studied. This led Yen and the writer[8] to assert that in most gases the negative ion had a greater mobility than the positive ion, but that where the mobilities were small the value of the negative mobility approached that of the positive mobility. This conclusion was also borne out by the work of Erikson[9,10] who found out that newly generated positive ions have the same mobility as the negative ions, but that after a short time they changed to the lower value of the mobility. This lead him to formulate a definite theory concerning the difference between the two ions. These measurements completely change the aspect of the problem and make an investigation of the ions in these gases from the point of view of the Erikson effect imperative. Such measurements are now in progress in this laboratory using the Erikson type of apparatus. That the negative mobility in HCl was less than the positive mobility was a conclusion arrived at in 1911 by Franck and Pringsheim[11] who studied the mobilities in a H_2—Cl_2 flame. They concluded from these

measurements that the positive ion in the flame was five times as mobile as the negative ion. They also stated that the free electrons did not exist in such flames and pointed out that all these results were a support to the cluster theory. As the conditions are markedly different in the flame from those in the present experiments it is not surprising that the results are somewhat different. They are, however, in general agreement with the results found here.

From the low negative mobility it also might be concluded that, at low pressures, the abnormal mobilities of the negative ions caused by finite paths travelled by these as electrons before attaching, observed in gases like air, would not appear. For a gas that depresses the negative mobility below that of the positive mobility would naturally be thought of as having a high electron affinity. Mobility measurements were made at 60 mm. and 16.3 mm. pressure on both ions. The former measurements were not particularly reliable as damp weather affected the electrometer on that day. The observed values of the mobilities reduced to 760 mm. assuming the inverse pressure law to hold are

PRESSURE	POSITIVE IONS	NEGATIVE IONS
60 mm	0.558	0.558
16.3 mm.	0.676	0.628

With the uncertainties involved at these low pressures the results constitute an adequate proof that the inverse pressure law holds. At the lower pressure the negative current voltage curve observed came sharply down to a zero value. It had no asymptotic foot nor did it have any indication of the dual nature of the curves observed with this apparatus by Wellisch[12] and Loeb[13] where electronic carriers are present. From this one may conclude that the electrons attach very rapidly to HCl molecules to make negative ions for the time of one alternation was about 0.002 second. Experimentally for the gases observed[14] the constant of attachment for electrons in this gas would be placed next to Cl_2, only in its low value.

Mobility Measurements in Mixtures of HCl and Air.—The results of mobility measurements made in mixtures of HCl and air are plotted in figure 1 as a function of the percentage of HCl gas present. The crosses represent the mobilities of the positive ions, the circled points the mobilities of the negative ions. The dotted curves represent mobilities of the positive and negative ions computed from the values of K_A and K_B for air and HCl from the composition of the mixture according to an equation which Blanc[1] found to hold for mixtures of H_2, CO_2 and air. The equation has the form

$$Kc = \frac{100\ K_A\ K_B}{(100 - c)\ K_B + CK_A}$$

where Kc is the mobility in the mixture containing c per cent of HCl. The full curves give the values computed according to an equation which the writer found to hold in the case of mixtures of NH_3 and air. This equation has the form

$$Kc = \frac{K_A K_B}{\sqrt{\frac{(100 - c)K^2_B + cK^2_A}{100}}}.$$

The values of K_A and K_B for pure air and pure HCl for positive and negative ions used in these computations were the observed values of 1.35 and 1.9 cm./sec. and 0.65 and 0.56 cm./sec., respectively. The lowest values for K_c for positive and negative ions at about 1% of HCl are not very reliable due to contact potential effects. The others are quite certain as prolonged pumping and subsequent measurements in air showed nearly normal mobilities in air. Thus, for example, in one case on 12 hours' pumping and measurement the mobilities rose from 1.175 and 1.59 cm./sec. for positive and negative ions to 1.26, and 1.86 cm./sec. for these ions. The prolonged pumping to get rid of the HCl present was necessitated by two conditions; first, the peculiar sensitivity of the mobilities to minute traces of HCl, and second, to the enormous absorption of HCl by the gold walls of the vessel and the gold leaf trap for mercury vapor. In fact, on first introducing HCl into the newly cleaned dry ionization chamber, an absorption of gas amounting to a cm. or so in pressure was observed in the course of the first two hours.

These results show at once that the negative mobilities fit the curve computed, according to the writer's method, exceptionally well from 100 per cent to 20 per cent HCl. The agreement between the positive mobility and the corresponding curve for the positive ions is not as satisfactory for the same range of composition. The deviation lies within the range of possible experimental error, while the excellent agreement for the negative ion is probably, in part, fortuitous. It is obvious that both the mobilities in this range agree with the equation holding for NH_3-air mixtures better than the equation applicable to the mixtures studied by Blanc. According to Debye[15] the applicability of the equation found for NH_3 and air can signify only one thing and that is that the ion changes the composition of the gas in which it moves by drawing more molecules of the reactive gas into its neighborhood.

Below concentrations of 20% HCl the points no longer fall on the curves at all. In fact the mobilities indicate that even traces of HCl gas change the mobility far more than would be expected. This can only be the case if the ions draw HCl molecules into their neighborhood thus increasing the concentration in their neighborhood. The effect is characteristic of what one would expect on the cluster ion theory. It was the effect looked for in NH_3 and air mixtures, but not found.

Thus the conclusion to be drawn from these measurements is that in HCl gas an actual cluster formation takes place, although the cluster, must be of the labile type consisting in an increased concentration of molecules of one type in the neighborhood of the ion. How this fact will agree with previous theoretical considerations and how the theory of ionic mobilities will be influenced by this new fact taken together with the experiments of Erikson, and Tyndall and Grindley will constitute the body of another paper.[15,16]

In conclusion, the writer wishes to acknowledge with thanks the timely assistanee of Mr. Wm. O. Smith in helping with some of the later measurements.

Summary.—Mobility measurements were made in pure dry HCl gas. The average mobility for the positive ion at 20°C. and 760 mm. was found to be 0.65 cm./sec. per volt/cm. and 0.56 cm./sec. per volt/cm. for the negative ion. This measurement of a negative mobility, definitely less than a positive mobility, constitutes the first exception to a general rule previously assumed from observation which indicated that the negative ion had a mobility either equal to or greater than the positive ion, *but not less*. The inverse pressure law of mobilities was found to hold for both positive and negative ions down to 16.3 mm. pressure. It was concluded from this that the constant of attachment of electrons to HCl molecules to form negative ions was nearly as low as that for Cl_2 gas. Finally, measurements of mobilities in mixtures of air and HCl showed unmistakable signs of a cluster ion formation in this gas, the mobilities being lowered by marked amounts for fractions of a per cent of HCl gas.

[1] Blanc, A., *J. Physique*, **7**, 825, 1908.
[2] Loeb and Ashley, *Proc. Nat. Acad. Sci.*, **10**, 351, 1924.
[3] Rutherford, E., *Phil. Mag.*, **44**, 422, 1897.
[4] Franck and Pohl, *Verh. D. Physik Ges.*, **9**, 69, 1907.
[5] Loeb, L. B., *J. Franklin Inst.*, **196**, 771, 1923.
[6] Wellisch, E. M., *Phil. Trans. Roy. Soc.*, **A209**, 249, 1909.
[7] Yen, K. L., *Proc. Nat. Acad. Sci.*, **4**, 106, 1918.
[8] Loeb, L. B., *Ibid.*, **11**, 428, 1925.
[9] Erikson, H. A., *Physic Rev.*, **20**, 111, 1922.
[10] Erikson, H. A., *Ibid.*, **24**, 502, 1924.
[11] Franck and Pringsheim, *Verh. D. Physik Ges.*, **13**, 328, 1911.
[12] Wellisch, E. M., *Phil. Mag.*, **34**, 55, 1917.
[13] Loeb, L. B., *Physic Rev.*, **17**, 89, 1921; *J. Franklin Inst.*, **197**, 45, 1924.
[14] Loeb, L. B., *Phil. Mag.*, **43**, 230, 1922.
[15] Loeb, L. B., *Proc. Nat. Acad. Sci.*, **12**, 42, 1926.
[16] Loeb, L. B., *J. Franklin Inst.* (to appear 1926).

THE MOBILITY OF GAS IONS IN HCl MIXTURES AND THE NATURE OF THE ION

BY L. B. LOEB

PHYSICAL LABORATORY, UNIVERSITY OF CALIFORNIA

Communicated December 7, 1925

The nature of the ordinary gaseous ion appears today to be as far from a definite solution as it was in 1909 when Wellisch[1] and Sutherland[2] independently put forward the small ion theory as a possible alternative to the then generally accepted cluster ion theory. While at different times it has been thought that various critical experiments decided the question in one way or another the writer[3,4] has shown that none of the evidence thus far has been conclusive. In a more recent paper on the theory of ionic mobilities the writer[4] concluded with J. J. Thomson[5] that if one assumed the mobilities to be controlled by the charge on the ion acting according to an inverse fifth power law, the actual nature of the ion had very little to do with the values of the mobilities as they were found in measurements in pure gases. The theory there advanced, while it agreed with most phenomena observed in such gases for ions better than previous theories, appeared to be seriously impaired by recent results of Tyndall and Grindley.[6] It also encountered some difficulty in explaining the recent results of Erikson.[7,8,9,10,11]

In 1923 Miss Ashley and the writer[12] pointed out that an ultimate answer to the question of the nature of the ion might be found through a study of the mobilities of the ions in mixtures of gases of widely different constitution. The answer to be derived from such a source would not, however, in any way contravert the theory already proposed, but would enable one to extend the knowledge of facts beyond the scope of the conclusions to be derived from that theory. To this end they investigated the mobilities of ions in mixtures of NH_3 and air. They found that for this case the mobility could be computed by an equation which *then* appeared to be in good agreement with the theory of the writer. This equation, as was stated above, seemed to be a consequence of assumptions which could not distinguish the presence of clustering did it occur. Hence while the agreement observed tended to confirm the theory it was not supposed to throw any light on the controversy. This equation found stated that the mobility K_c of an ion in a mixture of two gases A and B of masses M_A and M_B and of dielectric constants D_A and D_B could be computed from

$$K_c = \frac{K_A}{\sqrt{\frac{(100-C)+CQ}{100}}}$$

where c was the percentage of the gas B in the mixture, K_A the mobility of the ion in A, and Q was computed from the relation

$$Q = \frac{(D_B - 1)M_B}{(D_A - 1)M_A}.$$

This ratio, according to the theory, should be the same as the squared ratio of the mobilities K_A and K_B in the pure gases A and B. For NH_3 this was found to be the case within satisfactory limits. For other gases where the theoretical ratio Q does not agree as well with $(K_A/K_B)^2$ as in the case of NH_3 for both positive and negative ions the mobility in the mixture may be computed using $(K_A/K_B)^2$ in place of Q.

In studying mobilities of ions in mixtures of H_2, CO_2 and air, Blanc[13] in 1908 found that the mobility of the ion in the mixture could be computed with remarkable accuracy by an equation of the form

$$K'_c = \frac{K_A K_B}{\frac{(100 - c)K_B + CK_A}{100}}.$$

This equation does not fit the results in NH_3 and air mixtures nearly as well as an equation of the form used by the writer, to wit

$$K_c = \frac{K_A K_B}{\sqrt{\frac{(100 - c)K_B^2 + cK_A^2}{100}}}.$$

The essential difference is that the Blanc equation uses an average mobility weighted by the concentration, while the writer's equation uses the square root of the average squared mobilities weighted by the concentration. This difference between Blanc's equation and the writer's equation became evident to the writer some time after the papers mentioned were written and reflection showed that the theoretical mobility equation derived by the writer really requires the Blanc type of averaging instead of the one used by the writer. In discussing the theory of gas ion mobilities with Professor P. Debye the writer pointed out the observed difference in the equations holding for the two types of gaseous mixtures. Professor Debye ventured the opinion that the difference, if true, could only indicate one thing. That is, that in mixtures of the type NH_3 and air the ion changed the concentration of the gas in its neighborhood. Such an assumption is additional to those made in deducing the mobility equation above and definitely implies the existence of a labile cluster, in which the charged molecule exerts a choice on the molecules comprising the cluster. That is, to say that not only is there a cluster ion, but there is a preferential one in which the more active molecules predominate.

In view of the importance of this conclusion work was at once begun on a study of the mobilities in HCl and air mixtures to see whether the rule also held true there. The results of this investigation are given in a preceding paper[14] and seem in good accord with this conclusion. As shown in that paper, however, the results in the case of HCl are even more conclusive, for in HCl-air mixtures the initial lowering with traces of HCl is much greater than can follow from either method of averaging alone. It was the effect originally looked for in NH_3 but not found in that gas. It is possible that at very low concentrations the effect was present even in NH_3 but owing to complications introduced by contact potentials it was not detected. This indicates definitely the ability of the ion to gather around it a higher concentration of HCl than is present in the rest of the gas. That is, it indicates a definite labile cluster formation.

That such a change in composition of the gas in the immediate neighborhood of the ion must follow in the case of gases like NH_3 and HCl in air may be nicely shown from theoretical considerations. One may regard air and HCl molecules as being attracted to the ion with forces varying inversely as the fifth power of the distance. (Forces of this type are to be expected if a dielectric or conducting sphere is acted on by a radial field of force. They are also to be expected if molecular electrical dipoles which are constantly being oriented in any direction by their heat motions are acted on by a radial field of force. In both cases the induced dipole or the component of the fixed dipole in the direction of the field are acted on by a force varying inversely as the cube of the distance. The inducing force of the ionic charge on the dielectric sphere, or its orienting force on the dipoles, is the electric field produced at the point where the molecule finds itself. This varies inversely as the square of the distance. Hence the resultant attraction varies inversely as the fifth power of the distance.) Assuming such forces and using the dielectric constants for air and HCl gas, Mr. Edward Condon in this laboratory, in analogy with the law of atmospheres, has computed the ratio of the number of molecules of the two kinds of gas at different distances from the charged ion. His reasoning is as follows. Suppose a gas made up of two kinds of molecules 1 and 2. The density of the two kinds N_1 and N_2 is given in a field of force by the relation

$$N_1 = N_{10}\, e^{-\frac{E_1}{RT}} \qquad \text{and} \qquad N_2 = N_{20}\, e^{-\frac{E_2}{RT}}$$

where E_1 and E_2 are the potential energies of the molecules of each kind at the place where their relative densities are being calculated, and N_{10} and N_{20} are the densities at a point so far distant from the charge that the composition is that of the original mixture of the gases. If the molecules are acted on by forces that vary as $\frac{+K_1}{r^5}$ and $\frac{+K_2}{r^5}$ then the potential

energies at a distance r are $\frac{-K_1}{4r^4}$ and $\frac{-K_2}{4r^4}$ and one may write

$$\frac{N_2}{N_1} = \frac{N_{20}}{N_{10}} e^{\frac{K_2 - K_1}{4r^4 RT}}$$

where K_1 and K_2 are given by the following expressions,

$$K_1 = \frac{(D_1 - 1)\epsilon^2}{2\pi N_0}, \quad \text{and} \quad K_2 = \frac{(D_2 - 1)\epsilon^2}{2\pi N_0}.$$

In these expressions D_1 and D_2 are the dielectric constants of the gases 1 and 2, ϵ is the electron, N_0 is the Loschmidt's number, k is the Boltzmann constant, and T is the absolute temperature. Taking D_1 for air as 0.00059, and the recent value of Zahn[15] for D_2 for HCl at room temperatures as 0.0041, with N_2 as 2.7×10^{19}, the ratios of the number of ions for any initial mixture of HCl and air at any distance r from a singly charged ion can be found from the expression

$$\frac{N_2/N_1}{N_{20}/N_{10}} = e^{\frac{2.83 \times 10^{-29}}{r^4}}.$$

The values derived for a number of values of r comparable with ionic dimensions are given below:

R IN CM.	$\frac{N_2/N_1}{N_{20}/N_{10}}$
1.14×10^{-8}	10^{700}
4.0×10^{-8}	545
5.74×10^{-8}	13.5
6.9×10^{-8}	3.4
8.0×10^{-8}	2.34
9.2×10^{-8}	1.48
1.14×10^{-7}	1.17
1.38×10^{-7}	1.08
1.14×10^{-6}	1.000016
1.14×10^{-5}	1.00000000

It is thus seen that the molecules that find themselves at the surface of a cluster ion of 4.6×10^{-8} cm. radius are HCl molecules in the ratio of 546 to one if the mixture contains 50% HCl gas and air. In a mixture of 0.2% HCl and 99.8% air such a cluster would on the average consist of one HCl molecule to one air molecule. At greater distances from the ion the changes in concentration would be less striking and at the distance of one mean free path in a 50% HCl air mixture, or at 10^{-5} cm. there would virtually be no change in the concentration of the molecules due to the ion. In a homogeneous gas the mobilities of the ions of the Thomson[4,5] theory are chiefly influenced by those molecules which come close enough

to the ion to have their motion appreciably altered by its field of force. In a mixture, the molecules of higher dielectric constant, that is, those of greater activity on the ion, are as seen above present in far greater numbers in the immediate neighborhood of the ion than in the rest of the gas. Thus they are disproportionately active in altering the mobility of the ion in the gas. The effect of the increased concentration of the active molecules in the neighborhood of the ion is not included in the mobility equation deduced by the writer, for here the constitution of the gas was supposed to be uniform. In mixtures of gases of such widely divergent values of the dielectric constant it is not surprising that the mobility of the ions is lowered abnormally rapidly for small concentrations of the more active gas. On the other hand, it is to be expected that where the dielectric constants of the gases do not differ materially the mobilities will hardly be influenced by the changes of concentration. Thus it is not strange that Blanc found the simple law of mixtures to hold for the gases like H_2, air and CO_2 while in NH_3 and HCl and air mixtures the writer found an entirely different state of affairs. It would be expected that the effect of this change in concentration is more marked when traces of the gas are added. For since HCl is very active in lowering the mobility, and it was shown above that with only 0.2% of HCl a cluster would consist of an equal number of air and HCl molecules the initial lowering of the mobility would be much greater, than at concentrations of 50% HCl where the ion consisted almost exclusively of HCl molecules. That is, the presence of 50% of HCl molecules about the ion at low concentrations would be more effective than 100% at higher concentrations.

It would, at this juncture, be very difficult to deduce the expression for the mobility of the ions in such non-uniform mixtures of gases so that an attempt at quantitative verification would hardly be worthwhile. The qualitative reasoning outlined above shows that the peculiarities observed are to be expected in gaseous mixtures of widely differing dielectric constants on the writer's theory. They indicate that a definite "condensation" of molecules of higher dielectric constant about the nucleus of the ions occurs, even accompanied by a change in the concentrations of the constituents of the ion, so that it differs from the composition of the surrounding gas. Neither the theory nor these results can indicate how close or how permanent a union is effected between the ion and its attendant molecules. That the cluster of molecules closely bound to the charged molecule may be fairly stable will follow from the ratio of the potential energy of the ion to the kinetic energy of agitation. From the nature of the case it is not unlikely that in the course of the many collisions that an ion experiences in traversing a cm. of gas its cluster of attendant molecules may be changed many times. In general the net increase in density of all gaseous molecules in the neighborhood of an ion

might be considered as its cluster. The more distant of these molecules will constitute a very labile cluster indeed. Thus the investigation of the mobilities of ions in mixtures has been able to do what mobility measurements in pure gases could never do and that is to demonstrate the existence of a definite labile cluster ion.

The considerations of the nature of the cluster which have preceded this may with profit be carried further. The power of the ion to gather around it the more electrically active impurities present explains why such impurities as water vapor may, even in traces, profoundly influence the mobility.[16,17] The surrounding gas has nevertheless some influence and, in mixtures of air containing as little as 0.01% of HCl, the mobility of the ion would be largely that to be expected in pure air, for the ion would consist predominatingly of air molecules. This explains, in part, why the mobility of an ion in a gas takes on the value of the mobility of an ion in that gas no matter what the gas the ion was originally formed in. This was experimentally independently found first by Franck[16] and Wellisch and has later been very carefully investigated by Tyndall[6] and Grindley, and by Erikson.[8,10,11] That is, ions formed initially in CO_2 or H_2 and originally consisting of nearly exclusively CO_2 or H_2 molecules when measured while drifting through air have the same mobility as the ions generated initially in air. The effect described above is not, however, entirely ascribable to this effect. For the same authors found that even for ions consisting of the radioactive recoil atoms from Th D (whose mobility was studied by their radioactive effect so that it was certain that the initially ionized nucleus of the ion had retained the charge), showed about the same mobility (or even higher ones),* in air than the air ions formed in that gas showed. The measurements on this gas and such initially ionized molecules as H_2, $CHCl_3$ and CO_2 were accurate to 1%.

That the nucleus should play such an *exceedingly* minor rôle in the values of the ionic mobility is very strange indeed for it is contrary to all mobility equations based on dynamics. A detailed discussion of this will form part of a paper on ionic mobility equations to be published shortly. In conclusion the writer wishes to express his gratitude to Mr. Edward Condon for his help in formulating this discussion.

* These high mobilities were found in recent experiments of Erikson, very shortly after ionization. They may be due to doubly charged ions as the mobility is about twice the normal value.

[1] Wellisch, E. M., *Phil. Trans. Roy. Soc.*, **A209**, 249, 1909.
[2] Sutherland, W., *Phil. Mag.*, **18**, 341, 1909.
[3] Loeb, L. B., *J. Frank. Inst.*, **184**, 775, 1917.
[4] Loeb, L. B., *Phil. Mag.*, **48**, 446, 1924.
[5] Thomson, J. J., *Ibid.*, **47**, 342, 1924.
[6] Tyndall and Grindley, *Ibid.*, **47**, 689, 1924.
[7] Erikson, H. A., *Physic. Rev.*, **20**, 111, 1922.

[8] Erikson, H. A., *Physic. Rev.*, **26**, 629, 1925.
[9] Erikson, H. A., *Ibid.*, **24**, 502, 1924.
[10] Erikson, H. A., *Ibid.*, **24**, 622, 1924.
[11] Erikson, H. A., *Ibid.*, **26**, 465, 1925.
[12] Loeb and Ashley, *Proc. Nat. Acad. Sci.*, **10**, 351, 1924.
[13] Blanc, A., *J. Physique*, **7**, 825, 1908.
[14] Loeb, L. B., *Proc. Nat. Acad. Sci.*, **12**, 35, 1926.
[15] Zahn, C. T., *Physic. Rev.*, **24**, 400, 1924.
[16] Franck, J., *Jahrb. Rad. Elekt.*, **9**, 252, 1912.
[17] Wellisch, E. M., *Proc. Roy. Soc.*, **A82**, 500, 1909.

HIGH FREQUENCY RAYS OF COSMIC ORIGIN

By R. A. Millikan

Norman Bridge Laboratory, California Institute of Technolgoy

Read before the Academy November 9, 1925

It was as early as 1903 that the British physicists, McLennan and Burton[1] and Rutherford and Cooke[2] noticed that the rate of leakage of an electric charge from an electroscope within an air-tight metal chamber could be reduced as much as 30% by enclosing the chamber within a completely encircling metal shield or box with walls several centimeters thick. This meant that the loss of charge of the enclosed electroscope was not due to imperfectly insulating supports, but must rather be due to some highly penetrating rays, like the gamma rays of radium, which could pass through metal walls as much as a centimeter thick and ionize the gas inside.

In view of this property of passing through relatively thick metal walls in measurable quantity, the radiation thus investigated was called the "penetrating radiation" of the atmosphere, and was at first quite naturally attributed to radioactive materials in the earth or air, and this is in fact the origin of the greater part of it. But in 1910 and 1911 it was found that it did not decrease as rapidly with altitude as it should upon this hypothesis. The first significant report upon this point was made by the Swiss physicist, Gockel,[3] who took an enclosed electroscope up in a balloon with him to a height of 13,000 feet and reported that he found the "penetrating radiation" about as large at this altitude as at the earth's surface, and this despite the fact that according to Eve's[4] calculation it ought to have fallen to half its surface value in going up 250 feet.

In 1911, '12, '13 and '14 two physicists, Hess,[5] a Swiss, and Kolhörster,[6] a German, repeated these balloon-measurements of Gockel's, the latter going to a height of 9 km., or 5.6 miles, and reported that they found this

radiation decreasing a trifle for the first mile or so and then increasing until it reached a value at 9 km., according to Kolhörster's measurements, eight times as great as at the surface. This seemed to indicate that the penetrating rays came from outside the earth, and were, therefore, of some sort of cosmic origin. If so it was computed[7] that in order to fit the Hess and Kolhörster data the rays had to have an absorption coefficient of 0.57 per meter of water and an ionizing power within a closed vessel sent to the top of our atmosphere of at least 500 ions per cc. per sec., in place of the 10 or 12 ions found in ordinary electroscopes at the surface. The war put a stop the world over to further studies of this sort, but as soon as we could get the proper instruments built after the war in the newly equipped Norman Bridge Laboratory of Physics, I. S. Bowen and myself went to Kelly Field, near San Antonio, Texas, with four little recording electroscopes which we succeeded in the spring of 1922 in sending up in sounding balloons to almost twice the heights which had previously been attained. The highest flight reached the altitude of 15.5 km., or nearly 10 miles.

These instruments were interesting in that, though they were built of steel to hold 300 cc. of air at 150 lbs.' pressure, and were provided each with a recording barometer, thermometer and electroscope, also with two different sets of moving photographic films and the necessary driving mechanism, the total weight of the whole instrument was yet but 190 grams, or about 7 ounces. The altitudes were determined not only from the now well-established law of ascent of balloons, but also by direct, two theodolite observations which Major Wm. R. Blair of the U. S. Signal Corps kindly sent Lieutenant McNeil to Kelly Field for the express purpose of making for us.

In these experiments we expected, if the results previously reported were correct, to find very large rates of discharge; for our instruments went up to such heights that eighty-eight per cent of the atmosphere had been left beneath them, and only twelve per cent was left to cut down, by its absorption, the intensity of the hypothetical rays entering from outside. In other words, our electroscopes should have been exposed to radiations approaching in intensity those existing at the very top of our atmosphere. *We actually failed to find anything like the computed rates of discharge.* Our experiments were in agreement with those of the European observers in that our electroscopes showed a somewhat higher rate of discharge at high altitudes than at the surface, but at the same time they proved conclusively that *a radiation of the assumed properties did not exist, our observed rates of discharge being not more than one-fourth the computed amounts.*

Since the origin of the "penetrating rays" was still uncertain, Dr. Russell Otis and myself in the summer of 1923 went to the top of Pike's

Peak for the sake of making absorption experiments upon these radiation at the highest altitude to which we could carry large quantities of absorbing materials. For if the rays were not of cosmic origin they did not need to be more penetrating than are the gamma rays from radioactive materials, while if they were of cosmic origin the sounding balloon experiments of Bowen and myself had shown that they must be very much harder (more penetrating) than anybody had thus far assumed. What was needed was absorption experiments to determine just what sort of rays they actually were.

We carried 300 pounds of lead and a 6′ x 6′ x 6′ tank of water to the top of the peak and obtained as the net result of these absorption experiments the definite proof that the rays found at the top of Pike's Peak were predominantly of the hardness of ordinary gamma rays, and further that they were very largely, if not entirely, of local origin, since local conditions, such as a heavy snow storm and blizzard, which occurred while we were there, varied their intensity nearly as much inside a screen of 4.8 cm. of lead as outside. Kolhörster had by this time, after the brief publication of our Kelly Field data, and as a result, also, of new experiments made subsequently to them in crevasses and holes in glaciers in the Alps reduced his estimated absorption coefficients[8] from 0.57 to 0.25, a change he regards as within the limits of his experimental uncertainties, but a change which made the assumed rays so hard as to be no longer irreconcilable with our sounding balloon observations. But we found that our Pike's Peak observations were not yet compatible with his now (1923) assumed characteristics of rays of cosmic origin, viz., rays which produce 2 ions per sec. per cc. at the earth's surface, and have a coefficient of 0.25 per meter of water. For while in going from the altitude of Pasadena to that of Pike's Peak the number of ions observed with the unshielded electroscope increased from 11.6 to 22.2, an increase of 10.6 ions, the number of ions observed through the shield of 4.8 cm. of lead increased but from 9.37 to 11.6, an increase of only 2.23 ions. But radiation of the characteristics assumed above would have caused by itself, inside our lead screen, an increase of 3.34 ions, even if *none* of the large increase in radiation shown by the unshielded observations got through the lead shield—a supposition which we believed to be contrary to fact. In a word, our *Pike's Peak observations showed that if rays of cosmic origin existed at all they must be of different characteristics from any as yet suggested, and they further showed most interestingly that a very copious soft radiation of unknown origin existed at the altitude of Pike's Peak.*

Accordingly, Mr. Harvey Cameron and myself planned some new experiments for the summer of 1925 which were designed.

(1) To settle definitely the question of the existence or non-existence of a small, very penetrating radiation of cosmic origin—a radiation so hard

as to be uninfluenced by, and hence unobservable with the aid of, such screens as we had taken to Pike's Peak—and,

(2) To throw light on the cause of the variation with altitude of the radiation of gamma-ray hardness which our absorption experiments on Pike's Peak showed to be more than twice as copious there as at Pasadena.

The only possible absorbing material obtainable in the immense quantities needed, and of homogeneous and non-radioactive constitution, were the waters of very deep snow-fed lakes—snow-fed because the results of underwater experiments which we had previously carried on near Pasadena had been vitiated by our discovery that the waters were appreciably radioactive. We felt that there was much uncertainty as to how much this cause might have affected the European observations in and about glaciers. Further, our Pike's Peak experiments had demonstrated that if any of the penetrating rays were of *cosmic origin* the ionization due to them in our electroscope at sea level had to be much less than the 2 ions, assumed above, out of the 11.6 observed, the experimental error being, say, half an ion. No crucial tests could, therefore, possibly be made unless we could find very deep, non-radioactive lakes at very high altitudes where cosmic rays, if they existed, had two or three times the ionizing effect to be expected from them at sea level. We needed at least three ions due to cosmic rays, to vary with absorbing materials, if we were to obtain unambiguous evidence.

We chose for the first experiments Muir Lake (11,800 feet high), just under the brow of Mount Whitney, the highest peak in the United States, a beautiful snow-fed lake hundreds of feet deep and some 2000 feet in diameter. Here we worked for the last ten days in August, sinking our electroscopes to various depths down to 67 feet. *Our experiments brought to light altogether unambiguously a radiation of such extraordinary penetrating power that the electroscope-readings kept decreasing down to a depth of 50 feet below the surface.* The atmosphere above the lake was equivalent in absorbing power to 23 feet of water, so that here were rays so penetrating that, if they came from outside the atmosphere, they had the power of passing through $50 + 23 = 73$ feet of water, or the equivalent of 6 feet of lead, before being completely absorbed. The most penetrating X-rays that we produce in our hospitals cannot go through half an inch of lead. Here were rays at least a hundred times more penetrating than these, and having an absorption coefficient but one twenty-fifth, instead of "about one-tenth of that of the hardest known gamma rays."[8]

How unambiguous was the experimental evidence may be seen from the fact that with the aid of a new electroscope of high sensitivity the change in ions per cc. per sec. in going from the surface of Muir Lake to the depth of 15 meters (50 feet) was from 13.9 ions to 3.8 ions, or a decrease to about a fourth value. The largest decrease below a surface

reading reported by Kolhörster due to sinking electroscopes in water[8] was 2.1 ions, or a decrease of perhaps 20%, so that we have here obtained an altogether new precision of measurement and unambiguity of evidence.

To obtain definite evidence as to whether these very hard rays were of cosmic origin, coming in wholly from above and using the atmosphere merely as an absorbing blanket, we next went to another very deep snow-fed lake, Lake Arrowhead in the San Bernardino mountains, 300 miles farther south and 6700 feet lower in altitude, where the Arrowhead Development Company kindly put all their facilities at our disposal. The atmosphere between the two altitudes has an absorbing power equivalent to about 6 feet of water. *Within the limits of observational error, every reading in Arrowhead Lake corresponded to a reading 6 feet farther down in Muir Lake, thus showing that the rays do come in definitely from above, and that their origin is entirely outside the layer of atmosphere between the levels of the two lakes.*

Analysis of our absorption curves shows that the rays are not homogeneous but are hardened as they go through the atmosphere, just as X-rays are hardened by being filtered through a lead screen. Our hardest observed rays have an absorption coefficient of 0.18 per meter of water and the softest which get down to Muir Lake a coefficient of 0.3 per meter. The sounding balloon experiments of Bowen and myself make it improbable that they become very much softer than this at the top of the atmosphere, since otherwise we should have obtained larger readings in our very high flight.

Observations carried on day and night for four consecutive days on Pike's Peak at an altitude of 14,100 feet, and for two consecutive days on Mount Whitney at an altitude of 13,500 feet reveal no preferential direction in the heavens from which the rays come. *Within the limits of our uncertainty of measurement, then, these rays shoot through space equally in all directions.*

When absorption coefficients are reduced to wave length by a formula[9] of probable, though not yet of certain, validity our hardest observed rays have the wave-length 0.00038 Å, and those of longer wave-length go up to nearly twice this value, i.e., *we find a spectrum about an octave in width in a frequency region about 2000 times higher than that of the mean X-ray* (1 Å), *or as far above X-rays as X-rays are above light.* The shortest wave-length just computed corresponds to a frequency 10,000,000 times higher than that of visible light.

When these extraordinarily high frequency rays strike the earth, according to the now well-established Compton effect, *they should be transformed partially into soft rays of just about the hardness of the soft rays which we have actually observed on Pike's Peak and Mount Whitney.* The reason these soft rays were more plentiful on the mountain peaks than at Pasadena

would then be found simply in the fact that there are about three times as many of the hard rays to be transformed at the altitudes of the peaks as at that of Pasadena. This seems to be the solution of the second of our summer's problems.

We can draw some fairly reliable conclusions of a general sort as to the origin of these very penetrating and very high frequency rays. The most penetrating rays that we have known anything about thus far, the gamma rays of radium and thorium, are produced only by nuclear transformations within atoms. In other words, they are produced by the change of one atom over into another atom, or *by the creation of a new type of atom.* It is scarcely possible, then, to avoid the conclusion that these still more penetrating rays which we have here been studying are produced similarly by nuclear transformations of some sort. But these transformations must be enormously more energetic than are those taking place in any radioactive changes that we know anything about. For, according to our present knowledge, the frequency of any emitted ray is proportional to the energy of the subatomic change giving birth to it. *We can scarcely avoid the conclusion, then, that nuclear changes having an energy value perhaps fifty times as great as the energy changes involved in observed radioactive processes are taking place all through space, and that signals of these changes are being sent to us in these high frequency rays.*

The energy of the nuclear change that corresponds to the formation of helium out of hydrogen is known, and from it we have computed the corresponding frequency and found it to correspond closely to the highest frequency rays which we have observed this summer. *The computed frequencies of these cosmic rays also correspond closely to the energy involved in the simple capture of an electron by a positive nucleus.* Thus, the highest speed β-ray emitted by thorium leaves its mother atom with a speed which is equivalent to the energy acquired by the fall of an electron through 7,540,000[10] volts. This electron in order to get out of the mother atom was obliged to move against the pull upon it of the positive nucleus, and in this act it gained a potential energy the equivalent of a fall through 4,400,000 volts.[11] If this same electron had reversed its path and plunged into the nucleus it should have generated in so doing a 12,000,000 volt ray (7,540,000 + 4,400,000). The cosmic rays with which we have been dealing have frequencies which make them the equivalent of from 12 to 30 million-volt rays. It is not improbable that the capture of an electron by the nucleus of a light atom involves a higher energy than its capture by a heavy one, so that such captures as are here discussed constitute, perhaps, the most plausible hypothesis as to the origin of these rays.

Is it possible to imagine such a phenomenon going on all through space? The difficulty is not so insuperable, in view of the transparency even of large amounts of matter for these hard rays combined with Hubbell's

recent proof[12] at the Mount Wilson Observatory that some of the spiral nebulae are at least a million light years away. The centers at which these nuclear changes are taking place would then only have to occur at extraordinarily widely scattered intervals to produce the intensity of the radiation observed at Muir Lake.

The only alternative hypothesis to that above presented of high frequency rays traversing space in all directions, might seem to be to assume that the observed rays are generated in the upper layers of the atmosphere by electrons shooting through space in all directions with practically the speed of light. This hypothesis might help some in interpreting the mysterious fact of the maintenance of the earth's negative charge, but it meets with insuperable obstacles, I think, in explaining quantitatively the variation with altitude of the ionization in closed vessels. In any case, in its most important aspect, this hypothesis is very much like the one presented above, for it, too, fills space with rays of one sort or another traveling in all directions with the speed of light. From some such conception as this there now seems to be no escape. And yet it is a conception which is almost too powerful a stimulus to the imagination. Professor MacMillan of Chicago will wish to see in it evidence for the condensation into matter out somewhere in space of the light and heat continually being radiated into space by the sun and stars,[13] an altogether permissible speculation. Unfortunately the psychics will of course be explaining all kinds of telepathies with the aid of these cosmic rays. But, be that as it may, the simple experimental facts, as shown by the foregoing work, are:

(1) That these extraordinary penetrating rays exist;

(2) That their mass absorption coefficient may be as high as 0.18 per meter of water;

(3) That they are not homogeneous, but are distributed through a spectral region far up above X-ray frequencies—probably 1000 times the mean frequencies of X-rays;

(4) That these hard rays stimulate, upon striking matter, softer rays of about the hardness predicted by the theory of the Compton effect;

(5) That these rays come into the earth with equal intensity day and night and at all hours of the day or night, and with practically the same intensity in all directions.

Mr. I. S. Bowen, Dr. Russell Otis, Mr. G. Harvey Cameron and myself, all of whom have participated in this investigation and have received invaluable aid from the instrument maker, Mr. Julius Pearson, will publish full details of this work elsewhere.

[1] McLennan and Burton, *Physic. Rev.*, **16**, 184, 1903.
[2] Rutherford and Cooke, *Ibid.*, **16**, 183, 1903.
[3] Gockel, *Physik Zeit.*, **11**, 280, 1910.
[4] Eve, *Phil. Mag.*, **21**, 26, 1911.

[5] Hess, *Physik Zeit.*, **12**, 998, 1911, and **13**, 1084, 1912.
[6] Kolhörster, *Ibid.*, **14**, 1153, 1913, and *D. Physik Ges.*, July 30, 1914.
[7] E. v. Schweidler, Elster u.Geitel Fest schrift, p. 415, 1915.
[8] Kolhörster, *Sitz.-Ber. Preuss. Akad. Wiss*, **34**, 366, 23.
[9] N. Ahmad, *Proc. Roy. Soc.*, **A109**, 206, 1925.
[10] Report of Committee on X-Rays and Radioactivity of National Research Council, 1925, p. 92.
[11] Report of Committee on X-Rays and Radioactivity of National Research Council, 1925, p. 68.
[12] Hubbell, *Pop. Astron.*, **33**, pp. 252–255, 1925.
[13] MacMillan, *Science*, **62**, 122, 1925.

THE JOULE-THOMSON EFFECT IN AIR[1]

BY J. R. ROEBUCK

DEPARTMENT OF PHYSICS, UNIVERSITY OF WISCONSIN

Read before the Academy November 10, 1925

Air compressed to a chosen pressure, lying between 30 and 220 atm., was passed through a temperature equalizing coil of pipe in an oil bath held at one of a series of temperatures between 25° and 300°C., around a resistance thermometer bulb, through the wall of a round ended tube of porous procelain, around a second thermometer inside the porcelain tube, through a control valve where it dropped to atmospheric pressure and finally back to the compressor. The oil-bath temperature was held constant by means of electrical heating controlled by an electrical thermostat. The inlet pressure was held exceedingly constant by a rotating piston barostat which controlled a spill valve through which a small fraction of the compressed air could be wasted. The steadiness of the inlet pressure is vital to such work and a large amount of experimental attention was spent toward attaining it. The temperature of the inlet thermometer and the difference of temperature between the two thermometers was read on a Callendar-Griffiths bridge. The flow of air was measured roughly by a Venturi meter.

The performance of this form of porous plug was studied in detail, particularly as depending on the external and internal arrangements. Theory demands that the difference of temperature across the wall should depend only on the temperature and pressure conditions, and not at all on the rate of flow or particular plug. All the data reported satisfied these conditions, except at the low pressure (large volume) state where kinetic energy effects were very difficult to control.

To obtain the maximum amount of information, the measurements were arranged in groups, in which the pressure dropped from the same initial

pressure successively to a series of values of the lower pressure, giving thus a corresponding series of lower temperatures. A plot of these readings, T against p, gave a curve along which $(u + pv)$ remains constant—an isenthalpic curve. A group of such curves were obtained, one for each of the bath temperatures, 25, 50, 75, 100, 150, 200, 250, 283°C. In general, these curves show falling temperature with falling pressure, the latter being much greater at the lower temperatures and lower pressures.

The slope (μ) of these curves, called the Joule-Thomson coefficient, was calculated by taking the ratio of successive differences of temperature to the respective successive differences of pressure. These values of μ were plotted as a function of p and smooth curves (isenthalpics also) were drawn through them.

This data for μ was next arranged as isopiestics by picking the values of μ from the μ, p plot for a group of integral values of p and the corresponding values of T from the original p, T curves. Table 1 was read off from the curves so obtained, and gives the values of μ as a function of T and p.

TABLE 1

μ AS A FUNCTION OF T AND p

PRESS.	0°C.	25	50	75	100	150	200	250	280
1	0.2663	0.2269	0.1887	0.1581	0.1327	0.0927	0.0625	0.0402	0.0297
20	0.2494	0.2116	0.1777	0.1490	0.1244	0.0856	0.0564	0.0346	0.0246
60	0.2143	0.1815	0.1527	0.1275	0.1057	0.0708	0.0447	0.0251	0.0161
100	0.1782	0.1517	0.1283	0.1073	0.0890	0.0587	0.0347	0.0164	0.0078
140	0.1445	0.1237	0.1047	0.0875	0.0723	0.0467	0.0258	0.0093	0.0011
180	0.1125	0.0974	0.0833	0.0800	0.0578	0.0366	0.0185	0.0027	−0.0054
220	0.0812	0.0718	0.0627	0.0542	0.0452	0.0286	0.0127	−0.0020	−0.0110

The data shows that the value of μ falls steadily with rising temperature and with rising pressure till it goes through zero and becomes negative. The locus of $\mu = 0$ was read from these curves and also measured by special experiments. Table 2 read from the plot gives its course over the range covered by the experiments.

TABLE 2

P (ATM.)	T°C.
80	305
120	288
160	267
200	237

These values are widely different from previous experiments and calculations.

The temperature intercepts between the original isenthalpic curves can be used very directly to calculate the variation of specific heat (C_p) with pressure when C_p is known as a function of temperature. This latter was taken from the Reichsanstalt tables. The values so calculated are given in table 3.

TABLE 3

C AS A FUNCTION OF T AND p

PRESS.	0°C.	25	50	75	100	150	200	250	280
1	0.2405	0.2410	0.2415	0.2419	0.2424	0.2434	0.2443	0.2453	0.2458
20	0.2492	0.2487	0.2480	0.2475	0.2470	0.2466	0.2463	0.2468	0.2471
60	0.2656	0.2627	0.2603	0.2581	0.2562	0.2532	0.2512	0.2500	0.2492
100	0.2804	0.2760	0.2717	0.2681	0.2650	0.2602	0.2565	0.2536	0.2519
140		0.2873	0.2816	0.2767	0.2725	0.2658	0.2607	0.2566	0.2544
180		0.2960	0.2898	0.2840	0.2790	0.2707	0.2644	0.2596	0.2569
220		0.3020	0.2956	0.2893	0.2838	0.2748	0.2678	0.2622	0.2593

It will be observed that C_p increases with both pressure and temperature but at a decreasing rate.

The variation of C_p with pressure may be compared with the directly measured values. In table 4 the Reichsanstalt value for the average temperature 60°C. is given under C_{60}, and those of the present work under C, in both cases using the same value at 1 atm. as the starting point.

TABLE 4

COMPARISON OF C'S AT 60°C.

PRESS.	1	20	60	100	140	180	220
Factor	1.0	1.025	1.0725	1.118	1.157	1.189	1.2125
C	0.2419	0.2480	0.2595	0.2705	0.2799	0.2876	0.2933
C_{60}	0.2419	0.2473	0.2590	0.2699	0.2805	0.2884	0.2950
Diff. %	0	+0.28	+0.19	−0.22	−0.21	−0.28	−0.59

The value of μ may be used to calculate the correction to the air thermometer scale to reduce its readings to the thermodynamic scale. Using the available values for the coefficient of expansion between 0° and 100°C., leads to 273.15°K. as the temperature of the ice point on the thermodynamic scale.

The same general relation may be used to calculate the variation of this coefficient of expansion with pressure. In table 5 are given the values so calculated assuming the value at 1 atm. as given there. Under α_1 are Witkowski's directly measured data and under α_2 Holborn and Schulze's.

TABLE 5

COEFFICIENT OF EXPANSION

p	v_0	$I_p \times 10^4$	D	$E\%$	$\alpha \times 10^3$	$\alpha_1 \times 10^3$	$\alpha_2 \times 10^3$
0	∞		1.0		3.6610		
1	773.4	4.725	1.0025	0.003	3.6704		3.666
20	38.28	4.568	1.0503	0.05	3.845	3.83	3.826
60	12.54	3.869	1.1300	0.1	4.137	4.18	4.166
100	7.49	3.571	1.2009	0.2	4.397	4.41	4.424

The values for μ of the present paper are lower than Joule and Thomson's, and decidedly lower than Hoxton's. They are somewhat lower than Noell's, who covered about the same field but obtained very erratic data. This uncertainty may be attributed to lack of careful temperature

and pressure regulation as well as to very questionable thermal conditions about his plug.

The present work is to be extended as rapidly as possible to the region below room temperature.

[1] The full details will be found in *Proc. Amer. Acad., Boston*, **60**, 537, 1925.

ON NORMAL COÖRDINATES IN THE GEOMETRY OF PATHS

BY JOSEPH MILLER THOMAS*

PRINCETON UNIVERSITY

Communicated November 30, 1925

1. It has been shown that in terms of the normal coördinates for the geometry of paths,[1] the paths through the origin have linear equations and consequently are represented in an associated euclidean space by straight lines through the origin. It is well known that the paths are left unaltered by a change of the affine connection of the form

$$\Gamma^i_{jk} = \Gamma'^i_{jk} + \delta^i_j \varphi_k + \delta^i_k \varphi_j \qquad (1.1)$$

where φ_j is an arbitrary vector.[2] Under such a change the ordinary normal coördinates are not invariant,[3] but undergo some transformation of the group which leaves invariant the straight lines through the origin of the associated euclidean space. Professor Veblen[4] in his presidential address to the American Mathematical Society remarked that this transformation would be linear fractional if the vector φ were properly chosen. In the present paper is found a necessary and sufficient condition that the transformation of normal coördinates corresponding to a change of affine connection (1.1) be linear fractional at every point. Two other conditions which are sufficient, but not necessary, are also given.

As by-products we obtain a set of identities (2.15) connecting the components of the rth extension of a covariant vector with those of the covariant derivative of its $(r-1)$th extension.

2. Let y^i be the normal coördinates associated with a point $x^i = q^i$ and a coördinate system x. Let the corresponding equations of the paths be

$$\frac{d^2y^i}{ds^2} + C^i_{jk}\frac{dy^j}{ds}\frac{dy^k}{ds} = 0. \qquad (2.1)$$

If the affine connection be changed so that

$$C'^i_{jk} = C^i_{jk} - \delta^i_j \psi_k - \delta^i_k \psi_j,$$

the equations of the paths become[5]

$$\frac{d^2y^i}{dt^2} + C'^i_{jk}\frac{dy^j}{dt}\frac{dy^k}{dt} = 0.$$

Calling the normal coördinates associated with the C''s y'^i, we have the following differential equations for their determination[6]

$$\left(C'^\gamma_{jk}\frac{\partial y'^i}{\partial y^\gamma} - \frac{\partial^2 y'^i}{\partial y^j \partial y^k}\right)\frac{\partial y^j}{\partial y'^l}\frac{\partial y^k}{\partial y'^p}\, y'^l\, y'^p = 0. \qquad (2.2)$$

Moreover, from the initial conditions

$$y^i = 0, \quad \frac{\partial y^i}{\partial x^j} = \delta^i_j, \quad y'^i = 0, \quad \frac{\partial y'^i}{\partial x^j} = \delta^i_j, \quad \text{for } x^i = q^i,$$

it follows that the initial conditions for equations (2.2) are

$$y'^i = 0, \quad \frac{\partial y'^i}{\partial y^j} = \delta^i_j, \quad \text{for } y^i = 0. \qquad (2.3)$$

Hence the transformation of the associated euclidean space corresponding to the most general φ leaves the straight lines through the origin individually invariant.

We therefore only have to consider linear fractional transformations of the form

$$y'^i = \frac{y^i}{1 + a_j y^j} = \lambda y^i. \qquad (2.4)$$

Consequently we have

$$\frac{\partial y'^i}{\partial y^j} = \lambda(\delta^i_j - \lambda a_j y^i), \qquad \frac{\partial y^i}{\partial y'^j}\, y'^j = \frac{y^i}{\lambda},$$

$$\frac{\partial^2 y'^i}{\partial y^j \partial y^k} = -\lambda^2(\delta^i_j a_k + \delta^i_k a_j) + 2\lambda^3 a_j a_k y^i.$$

Substituting these results in equations (2.2), we find on using the relations[7]

$$C^i_{jk} y^j y^k = 0,$$

that a necessary and sufficient condition for linear fractional transformation of the normal coördinates is

$$(\psi_i - \lambda a_i)y^i = 0. \qquad (2.5)$$

Suppose that ψ_i is given by the series

$$\psi_i = b_i + a_{ij}y^j + \tfrac{1}{2!}\, a_{ijk}y^j y^k + \ldots. \qquad (2.6)$$

Substituting from (2.6) in (2.5) and equating coefficients to zero, we obtain

$$b_i = \psi_i(0, 0, \ldots, 0) = a_i, \tag{2.7}$$

$$\frac{\partial \psi_i}{\partial y^j} + \frac{\partial \psi_j}{\partial y^i} + 2\psi_i\psi_j = 0, \tag{2.8}$$

$$P\left(\frac{\partial^{r-1}\psi_i}{\partial y^j \partial y^k \ldots \partial y^l}\right) + (-1)^r \underline{|r}\ \psi_i\psi_j \ldots \psi_l = 0, \quad (r = 2, 3, \ldots) \tag{2.9}$$

where the symbol P denotes the sum of the terms obtained from those within the parentheses by cyclic permutation of the subscripts $i, j, \ldots, l$, and where all of the expressions are evaluated at the origin of the normal coördinates.

Conditions (2.7) can always be satisfied by proper choice of the constants a_i. The conditions (2.8) and (2.9) can be written

$$\varphi_{i,j} + \varphi_{j,i} + 2\varphi_i\varphi_j = 0, \tag{2.10}$$

$$P(\varphi_{i,jk\ldots l}) + (-1)^r \underline{|r}\ \varphi_i\varphi_j \ldots \varphi_l = 0, \tag{2.11}$$

where φ_i are the components of ψ in the x coördinate system, and where $\varphi_{i,jk\ldots l}$ denotes the $(r-1)$th extension[8] of φ_i formed with respect to the affine connection Γ^i_{jk}. Equations (2.10) and (2.11) must hold at all points of the x manifold in order that the normal coördinates associated with every point may undergo a linear fractional transformation.

Suppose that equations (2.10) are satisfied. Forming the first extension of them, we get

$$\varphi_{i,j,k} + \varphi_{j,i,k} + 2\varphi_{i,k}\,\varphi_j + 2\varphi_i\,\varphi_{j,k} = 0.$$

These equations can be written[9]

$$\varphi_{i,jk} + \varphi_{j,ik} - 2\varphi_\alpha A^\alpha_{ijk} + 2\varphi_{i,k}\varphi_j + 2\varphi_i\varphi_{j,k} = 0.$$

Permuting the subscripts cyclically, adding and using the well known identity

$$P(A^\alpha_{ijk}) = 0$$

and the equations (2.10), we find

$$P(\varphi_{i,jk}) - \underline{|3}\ \varphi_i\varphi_j\varphi_k = 0.$$

By induction we can show that equations (2.11) for $r > 2$ are consequences of (2.10). To this end we assume (2.11) hold for r subscripts. Forming the covariant derivative of these equations, we get

$$P[\varphi_{i,jk\ldots l,p} + (-1)^r \underline{|r}\ \varphi_{i,p}\varphi_j \ldots \varphi_l] = 0,$$

where it must be remembered that P refers to the cyclic permutation of

$i,j,\ldots,l$ only. If we now permute all the subscripts cyclically and add the results, we find

$$P'[\varphi_{i,jk\ldots l,p} + (-1)^r \underline{|r} \overset{*}{\varphi}_{i,p}\varphi_j\ldots\varphi_l] = 0,$$

where the symbol P' denotes cyclic permutation of the indices $i,j,\ldots,l$ followed by cyclic permutation of all the indices $i,j,\ldots,p$. It is readily seen that the expression arising from the second term in the brackets is symmetric in every pair of the subscripts, so that $(\varphi_{i,p} + \varphi_{p,i})/2$ can be written in place of $\varphi_{i,p}$. Doing this and making use of (2.10), we get

$$P'(\varphi_{i,jk\ldots l,p}) + (-1)^{r+1}\, r|\underline{r+1}\ \varphi_i\varphi_j\ldots\varphi_p = 0. \qquad (2.12)$$

From the vector property of φ it follows that

$$\psi_i = \varphi_\alpha \frac{\partial x^\alpha}{\partial y^i}.$$

Differentiation and evaluation at the origin of normal coördinates give

$$\left(\frac{\partial^{r-1}\psi_i}{\partial y^j\partial y^k\ldots\partial y^l}\right)_0 = \varphi_{i,jk\ldots l} = \frac{\partial^{r-1}\varphi_i}{\partial x^j\partial x^k\ldots\partial x^l} + \ldots, \qquad (2.13)$$

where the unwritten terms form an expression linear and homogeneous in φ_i and their derivatives. The coefficients are rational integral expressions in the functions[10] $\Gamma^\alpha_{\beta\gamma\ldots\delta}$.

Differentiation of (2.13) gives

$$\frac{\partial\varphi_{i,jk\ldots l}}{\partial x^p} = \frac{\partial^r\varphi_i}{\partial x^j\partial x^k\ldots\partial x^p} + \ldots,$$

the coefficients of φ_i and its derivatives now involving the derivatives of $\Gamma^\alpha_{\beta\gamma\ldots\delta}$. If in these general formulas we interpret x^i as normal coördinates and evaluate at the origin, the derivatives of φ_i become its extensions, the quantities $\Gamma^\alpha_{\beta\gamma\ldots\delta}$ reduce to zero and their derivatives become the generalized normal tensors[11] $A^\alpha_{(\beta\ldots\gamma)\delta\ldots\epsilon}$.

Hence, we have

$$\varphi_{i,jk\ldots l,p} = \varphi_{i,jk\ldots lp} + \ldots, \qquad (2.14)$$

the unwritten terms forming an expression linear and homogeneous in φ_i and their derivatives, each coefficient being a generalized normal tensor.

Now subject the subscripts in (2.14) to all the permutations in the symmetric group on $(r+1)$ letters, and add the resulting equations. The coefficients of any extension of φ_i, other than those written explicitly in (2.14), will consist of the sum of generalized normal tensors of the type $A^\alpha_{(\beta\ldots\gamma)\delta\ldots\epsilon}$, the subscripts occurring once and only once in each of the possible orders. The coefficients are therefore zero.[12] The only terms remaining arise from those written explicitly in (2.14). Since the first

of these tensors is symmetric in $j,k,\ldots,l$ and the second in $j,k,\ldots,l,p$, we find on dividing by $\underline{|r-1}$

$$P'(\varphi_{i,}{}^{*}{}_{jk\ldots l,p}) = rP(\varphi_{i,jk\ldots lp}), \tag{2.15}$$

the P now referring to the cyclic permutation of all the subscripts.

It is to be noted that the formulas (2.15) are identities satisfied by the extensions of any covariant vector.

Substitution from (2.15) in (2.12) completes the induction. Hence equations (2.10) constitute sufficient as well as necessary conditions.

THEOREM 1. *A projective change of affine connection brings about a linear fractional transformation of the normal coördinates at every point if, and only if, the vector φ defining the change of connection satisfies the relations*

$$\varphi_{i,j} + \varphi_{j,i} + 2\varphi_i\varphi_j = 0.$$

3. It is known[13] that the Ricci tensors for two affine connections related by (1.1) are connected by the following equations:

$$R'_{ij} = R_{ij} + n\varphi_{i,j} - \varphi_{j,} + (n-1)\varphi_i\varphi_j. \tag{3.1}$$

Interchanging i and j and adding, we find

$$R'_{ij} + R'_{ji} = R_{ij} + R_{ji} + (n-1)(\varphi_{i,j} + \varphi_{j,i} + 2\varphi_i\varphi_j).$$

Hence conditions (2.10) are equivalent to

$$R'_{ij} + R'_{ji} = R_{ij} + R_{ji}, \qquad (n > 1).$$

We can state this result in the following

THEOREM 2. *In order that the normal coördinates at every point undergo a linear fractional transformation when the affine connection is changed so as to leave the paths invariant, it is necessary and sufficient that the symmetric part of the Ricci tensor be preserved.*

4. From (3.1) it is seen that a necessary and sufficient condition that the Ricci tensor be preserved *in toto* is

$$\Phi_{ij} \equiv \varphi_{i,j} + \varphi_i\varphi_j = 0. \tag{4.1}$$

Since conditions (2.10) are satisfied if (4.1) are, the preservation of the Ricci tensor is a sufficient condition.

The curvature tensors[14] for the two connections are related as follows:

$$B'^{i}_{jkl} = B^{i}_{jkl} + \delta^{i}_{j}(\Phi_{lk} - \Phi_{kl}) - \delta^{i}_{k}\Phi_{jl} + \delta^{i}_{l}\Phi_{jk}.$$

From these last relations it readily follows that

$$B'^{i}_{jkl} = B^{i}_{jkl}$$

if, and only if, conditions (4.1) are satisfied. The preservation of the curvature tensor is therefore a second sufficient condition.

THEOREM 3. *A projective change of affine connection preserve the Ricci tensor if, and only if, it preserves the curvature tensor.*[15] *Such a change brings*

about a linear fractional transformation of the normal coördinates at every point.

* NATIONAL RESEARCH FELLOW IN MATHEMATICS.

[1] Cf. O. Veblen, these PROCEEDINGS, **8**, 1922, p. 192.

[2] H. Weyl, *Göttinger Nachrichten*, 1921, p. 99.

[3] A type of normal coördinates invariant under the change (1.1) has been given by O. Veblen and J. M. Thomas in these PROCEEDINGS, **11**, 1925, p. 204.

[4] O. Veblen, "Remarks on the Foundations of Geometry," *Bull. Amer. Math. Soc.*, **31** (1925), p. 131.

[5] Cf. O. Veblen and T. Y. Thomas, "The Geometry of Paths," *Trans. Amer. Math. Soc.*, **25**, 1923, p. 557. This paper will be referred to as G.

[6] G., p. 563.

[7] G., p. 562.

[8] G., p. 571.

[9] G., p. 575, formula (12.1).

[10] G., p. 560.

[11] G., p. 577.

[12] G., p. 579, formula (13.16).

[13] G., p. 559, formula (5.10).

[14] G., p. 559.

[15] In an article which appeared after the present paper had been sent to the printer, J. A. Schouten gives the result contained in the first sentence of this theorem. Cf. *Trans. Amer. Math. Soc.*, **27**, 1925, p. 453.

A SIMPLE DERIVATION OF KRONECKER'S RELATION AMONG THE MINORS OF A SYMMETRIC DETERMINANT

BY E. B. STOUFFER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF KANSAS

Communicated November 19, 1925

In 1882 Kronecker[1] stated without proof that he had discovered the following theorem:

Among the minors of order m of a symmetric system

$$a_{ik} = a_{ki}$$

there exists the linear relation

$$(A) \qquad |a_{gh}| = \sum_r |a_{ik}|$$

$(g = 1, 2, \ldots, m;\ h = m + 1, \ldots, 2m)$

$(i = 1, 2, \ldots, m - 1, r;\ k = m + 1, \ldots, r - 1,\ m, r + 1, \ldots, 2m;\ r = m + 1, m + 2, \ldots, 2m).$

Numerous proofs and extensions[2] of this important relation have been given since its first announcement. It is the purpose of the present

note to obtain Kronecker's relation in a very simple manner by the use of certain operators.

Let us first observe that the double operation

$$\sum_{j=m}^{2m} a_{rj} \frac{\partial}{\partial a_{mj}} \left(\sum_{l=1}^{m} a_{lm} \frac{\partial}{\partial a_{lr}} |a_{gh}| \right)$$

first replaces the elements of column r of $|a_{gh}|$ by elements with the same first subscript, but with the second subscript m and then replaces the elements of the last row of this new determinant by elements with the same second subscript but with the first subscript r. Consequently,

$$\sum_{r=m+1}^{2m} \sum_{j=m}^{2m} a_{rj} \frac{\partial}{\partial a_{mj}} \sum_{l=1}^{m} a_{lm} \frac{\partial}{\partial a_{lr}} |a_{gh}| = \sum_{r} |a_{ik}|, \qquad (1)$$

where $\sum_{r} |a_{ik}|$ is exactly the right member of relation (A).

The left member of (1) is equivalent to

$$\sum_{r=m+1}^{2m} \sum_{j=m+1}^{2m} a_{rj} \left[\sum_{l=1}^{m} a_{lm} \frac{\partial}{\partial a_{mj}} \frac{\partial}{\partial a_{lr}} |a_{gh}| \right] + \sum_{r=m+1}^{2m} a_{rm} \frac{\partial}{\partial a_{mr}} |a_{gh}|. \quad (2)$$

Since $a_{rm} = a_{mr}$, we have

$$\sum_{r=m+1}^{2m} a_{rm} \frac{\partial}{\partial a_{mr}} |a_{gh}| = |a_{gh}|.$$

The element $a_{rj} = a_{jr}$ $(r > m, j > m)$ occurs linearly and homogeneously in the first part of (2). The coefficient of $a_{rr} = a_{jj}$ vanishes because of the two successive partial derivatives with respect to elements of the same column of $|a_{gh}|$. The coefficient of $a_{rj} = a_{jr}$ $(r \neq j)$ is

$$\sum_{l=1}^{m} a_{lm} \left(\frac{\partial}{\partial a_{mj}} \frac{\partial}{\partial a_{lr}} |a_{gh}| + \frac{\partial}{\partial a_{mr}} \frac{\partial}{\partial a_{lj}} |a_{gh}| \right).$$

The expression in parenthesis vanishes, a fact which becomes evident if we interchange columns r and j in one of the determinants.

The left side of (1) is thus reduced to $|a_{gh}|$ and we have the relation (A).

Many other relations involving series of determinants may be obtained by the use of operators similar to those employed above. Some of the relations will be given in future papers.

[1] Kronecker, L., *Werke*, **2**, (391–396).

[2] Runge, C., *J. für Math.*, **93**, 1882 (319–327).

Mehmke, R., *Math. Ann.*, **26**, 1885 (209–210).

Schendel, L., *Zeit. Math. Phys.*, **32**, 1885 (119–120).

Muir, T., *Proceed. R. Soc. Edinburgh.*, **15**, 1888 (96–105); *Ibid.*, **23**, 1900 (142–154).

White, H. S., *Amer. Math. Soc. Bull.*, **2**, 1896 (136–138).

Metzler, W. H., *Amer. Math. Soc. Trans.*, **2**, 1901 (395–403).

Nanson, E. J., *Messenger of Math.* (Ser. 2) **31**, 1901 (140–143); *Amer. J. Math.*, **27**, 1905 (69–76).